ADAPTIVE DYNAMICS
THEORETICAL BEHAVIORISM

Orientation, Reflexes, Habituation, Feeding, Search and Time Discrimination

J. E. R. Staddon

(MIT-Bradford, 2001, slightly revised, and corrected, 2019)
In Memory of my Father,

an Ingenious Man
ADAPTIVE DYNAMICS

J. E. R. Staddon

Preface

Chapter 1: THEORETICAL BEHAVIORISM: AIM and METHODS

Chapter 2: ADAPTIVE FUNCTION, I: THE ALLOCATION OF BEHAVIOR

Chapter 3: ADAPTIVE FUNCTION, II: BEHAVIORAL ECONOMICS

Chapter 4: TRIAL-AND-ERROR

Chapter 5: REFLEXES

Chapter 6: HABITUATION and MEMORY DYNAMICS

Chapter 7: MOTIVATION, I: FEEDING DYNAMICS and HOMEOSTASIS

Chapter 8: MOTIVATION, II: A MODEL FOR FEEDING DYNAMICS

Chapter 9: MOTIVATION, III: FEEDING and INCENTIVE

Chapter 10: ASSIGNMENT of CREDIT

Chapter 11: STIMULUS CONTROL

Chapter 12: SPATIAL SEARCH

Chapter 13: TIME, I: PACEMAKER-ACCUMULATOR THEORIES

Chapter 14: TIME, II: MULTIPLE-TIME-SCALE THEORY

Chapter 15: TIME, III: MTS and TIME ESTIMATION

Chapter 16: AFTERWORD

References
PREFACE

Mach’s weakness, as I see it, lies in the fact that he believed more or less strongly that science consists merely of putting experimental results in order; that is, he did not recognize the free constructive element… He thought that somehow theories arise by means of discovery and not by means of invention. Albert Einstein

This book is an argument for a simple proposition: That the way to understand the laws and causes of learning in animals and man is through the invention, comparison, testing, and modification or rejection of parsimonious black-box models. The models I discuss are neither physiological nor cognitive. They are behavioristic, in the most elementary sense: “One interesting definition [of the concept of thinking] has been proposed by A. M. Turing: a machine is termed capable of thinking if it can…imitate a human being by answering questions sufficiently well to deceive a human questioner for a reasonable period of time. A definition of this type has the advantage of being operational or, in the psychologists’ term, behavioristic.” (Shannon & McCarthy, 1956, p. v) Turing wanted to compare human and machine verbal protocols; I will be comparing the real-time behavior of animal and machine.

Behaviorism is unfashionable these days. It lost favor not, I believe, because it is fundamentally wrong, but because of the biases of its gurus. J. B. Watson eschewed internal states entirely. Clark Hull, now almost forgotten but once hugely influential, thought that a theory of learning could be built on fragmentary understanding of the nervous system, using the conceptual tools of philosophical positivism. But positivism (like all philosophy of science) is descriptive, not prescriptive. Hull’s formal skills were modest; the computational resources available to him were primitive; and his between-group experimental method is ill-suited to the discovery of learning processes in individual organisms. It is no wonder, therefore, that his ambitious attempt largely failed. More successful was the radical behaviorism of B. F. Skinner, because his experimental methods were more powerful, his writing style more persuasive and his public relations—linking operant conditioning to the universal cause of human betterment—much more effective. But Skinner also failed, because he did not understand or value theory. He condemned theories that involve “explanation of an observed fact which appeals to events taking place somewhere else, at some other level of observation, described in different terms, and measured, if at all, in different dimensions” (Skinner, 1950, p. 193)—an odd view that rules out of science the gene theory, atomic theory and indeed almost any kind of reductionism. Since even the most elementary learning process, habituation, requires at least one hidden (“internal-state”) variable for an adequate description, Skinner’s proscription of hidden variables effectively ruled out any real learning theory at all (Staddon, 1997a). But theory-phobia is an historical accident, not intrinsic to behaviorism.

Reflecting its ancestry in experimental psychology, the psychology of learning has for too long regarded the collection of data as its main task. Such theories as there are—most obviously Hullian theory, but also more recent theories of choice, classical conditioning and interval timing—are like elaborate Renaissance oil paintings: constructed incrementally, brush stroke by brush stroke, experiment by experiment. They are built with the confidence that their foundations are secure. And all tacitly accept physiological reductionism, the idea that elements of the perfected theory should map easily onto neurophysiology. The addition of neural networks to the theoretical toolkit in the 1970s only underlined this implicit commitment to psychoneural isomorphism—despite the vast gulf between the simple logical structure of back-propagation networks and the messy confusion of real neurophysiology.

The alternative I offer is also theoretical, but more like watercolor than oil painting. Rather than building up a massive structure step by step, I propose instead a succession of simple sketches, each designed to capture some aspect of the learning process. I argue that our understanding of behavior can advance only through what Einstein called the “free constructive ele-
The invention and testing of formal models for behavior. I emphatically dispute an unstated premise of the purely empirical approach Einstein condemned: “that science consists merely of putting experimental results in order”. Conjecture, not just about variables, but about processes, is essential.

The models I propose are simple, physiologically agnostic and disposable. Parsimony is a primary value. The data are entirely behavioral. If the model nevertheless seems to fit well with physiology (as it does, in a few cases) so much the better. But the aim is to explain behavior, not brain-behavior relations. And as each model is proposed and, eventually, refuted, we may hope to learn a little more about the essential mechanisms of learning—so the next model will be a little better, a bit more securely founded. Some history, and the problems with bottom-up physiological reductionism, are discussed in Chapter 1.

The central fact about behavior is that it is a process, a series of events that occur in time. It is dynamic. So, my main emphasis is on the way behavior changes in time. As to the important “variables of which behavior is a function”, in Skinner’s phrase, I believe that the best guide is function, but in Tinbergen’s (1963) sense of “adaptive function”—studied formally as optimality analysis—not Skinner’s. Optimality analysis, which answers the question, “What does this behavior maximize?” is perhaps the most successful attempt to date to provide a unified account of operant (instrumental) behavior in animals. Chapters 2 and 3 summarize psychological and economic optimality accounts of operant behavior. I conclude that economics is not fundamental, but it can point to variables that are likely to be important in causal accounts. Chapter 4 serves two functions: it is an introduction to elementary dynamics; and it offers models of the simplest kind of adaptive behavior, trial-and-error, in some very simple organisms. This chapter introduces most of the theoretical ingredients I use in later chapters. Chapter 5 shows how the properties of reflexes, described in every introductory learning text, can be derived from simple dynamic elements. Chapter 6 introduces a major topic: habituation, the most elementary learning process and one linked (I will argue) to general principles of memory and interval timing. The next three chapters, 7, 8 and 9, discuss feeding regulation, a process involved in almost every reinforcement-learning experiment, but one whose dynamics are still not fully understood. I offer an extremely simple model for the regulatory dynamics of feeding and show the model can explain how rats adapt to a wide range of environmental feeding challenges. Chapter 10 discusses assignment of credit, an essential function of operant (What did I do that caused that?) and classical (What stimulus signaled that?) conditioning. The chapter ends with questions rather than answers, because I offer no model for classical conditioning or associative processes in general. Chapters 11 and 12 nibble at the edges of the associative question, discussing a very simple dynamic model for stimulus generalization and applying it to spatial navigation. Chapters 13, 14 and 15 deal with another associative question: interval timing. I conclude that there may be no dedicated “interval clock.” Instead interval timing, habituation and human forgetting may all rely on the same memory process. The final chapter summarizes the basic approach. Chapters are roughly progressive, in the sense that earlier chapters assume less subject knowledge and mathematical sophistication than later ones.

John Staddon
Durham, North Carolina
ACKNOWLEDGEMENTS

I thank the members of the Learning and Adaptive Behavior Group at Duke for many stimulating discussions over the years. I thank especially my longtime collaborator Jennifer Higa, whose enthusiasm and willingness to entertain half-baked ideas allowed some of them to become fully baked. Nancy Innis has also provided encouragement and historical insights over many years. Richard Palmer contributed equally expertise and skepticism to our discussions and conveyed some measure of mathematical ‘taste’. I thank Silvano Zanutto, with whom I worked on the feeding model discussed in Chapters 7-9. Alliston Reid is responsible for much of the work on generalization and spatial navigation discussed in Chapters 11 and 12, to which Mircea Chelaru also made important contributions. Chelaru and Clive Wynne collaborated on some of the interval timing work. My students (now colleagues) Armando Machado and Valentin Dragoi contributed both criticism and insight. I expect they will improve on what is offered here. I also gratefully acknowledge a stimulating longtime association with Juan Delius and, more recently, collaborative assistance and criticism from Dan Cerutti. Finally, I acknowledge early inspiration from David McFarland, whose pioneering theoretical work in the 1960s and 1970s on feedback mechanisms in animal behavior has not received the recognition it deserves.

The late Harry Stanton, of Bradford books, provided early, and apt criticism, as well as the initial impetus for the book. My assistants Mary Michael and, most recently, Elizabeth Mayhall, coped good-humoredly with sometimes erratic demands, unexpected deadlines and the steady drone of bureaucratic imperatives.

As always, without my wife’s support, little would have been accomplished.

I gratefully acknowledge support from the Alexander von Humboldt Foundation, NIMH, NSF, NIDA and the Military-University Research Initiative (MURI) of the U. S. Army. Without a Senior Research Scientist Award from NIMH, and Duke University’s willingness to cooperate by relieving me of most teaching and administrative responsibilities, this work could never have been done.
Chapter 1

THEORETICAL BEHAVIORISM: AIM AND METHODS

Simple isn’t the same as obvious. Many years ago, I attended a lecture by Georg von Békésy, distinguished physiologist and experimental psychologist and, a few years later, Nobel Laureate for his work on the mechanisms of hearing. Békésy was lecturing to the large introductory course in natural science at Harvard. The course was organized along “parade of stars” lines, with various luminaries giving elementary summaries of their special topics in one or two lectures each. Békésy spoke in the vast bowl-like Allston Burr auditorium, freshly built in the brutal concrete then fashionable (but now demolished, perhaps because too many students plummeted from its higher tiers). I was sitting behind an undergraduate and because of the steep rake of the seats, I could see her notepad clearly. She began the lecture by writing furiously. But as elegant experiment followed elegant experiment—each one introduced with the phrase (in Békésy's thick Hungarian accent) “It is very simple to see...” her note-taking began to flag. After thirty minutes or so she turned to a fresh page and scrawled in large, block capitals “IT IS VERY SIMPLE...”

Békésy’s work is simple, in the scientific sense44, but obviously the ability to discern scientific simplicity is not inborn. Often forgotten in psychology is the great truth that in science, simplicity—of a more subtle kind, perhaps, than can be immediately appreciated by a beginning undergraduate—is the royal road to scientific truth. Not that every simple theory is true; experimental test after all trumps simplicity. But when two theories are compared, the simpler, in a special sense I will try and convey, is almost invariably the better one. This book aims to present a picture of the psychology of learning that is simple in Békésy's sense.

I take for granted two things with which no biobehavioral scientist would quarrel: That organisms are machines; and that the physical machinery for behavior is biological—nerves, muscles and glands. But I also argue for a third thesis, not universally accepted: that an essential step in understanding how all this machinery works is dynamic black-box modeling. I propose that a black-box model will always provide the most compact summary of behavior. And such an account is likely to be much simpler than a neurophysiological account, if indeed one can be found (we are a long way from understanding the neurophysiological underpinnings for most behavior). I also propose that an accurate model may be essential to discovering the physical processes that underlie behavior, because it tells the neuroscientist just what task the brain is carrying out. Since this view differs in some respects from cognitive psychology, which also deals in models, and since its heritage is behavioristic, I call it theoretical behaviorism (Staddon, 1993b,c, 1998; see Staddon, in press, for a fuller account of theoretical behaviorism).

This chapter sets the framework for what follows. First, I discuss the relation between behavioral studies — theory and experiment — and neurophysiology. Many behavioral scientists, and almost all neuroscientists, still think that psychology (the study of behavior) is in principle subordinate to neuroscience (the study of the brain). They are wrong, and I’ll try and explain why. If the secrets of behavior are not to be found by expanding our understanding of neurons and brain function, how is the job to be done? Through experiment and the creation and testing of simple models, I believe. The latter part of the chapter discusses the role of models and explains the two types of experiments that are used to study behavior. The rest of the book will explore some models that my colleagues and I have found useful in the study of learning.

44 For some samples see his wonderful little book Sensory Inhibition (1967).
**Brain and Behavior**

*Evolution in Silicon.* The best way to explain, and defend, the thesis that modeling is at the core of psychology is through an example not from psychology but from engineering. The constructions of engineers are not intrinsically mysterious; and we all know, or think we know, how machines work. If I can show, therefore, that the physical workings of some modern machines are almost as hard to understand as the neurobiology of behavior, the approach to psychobiology that I propose will seem less shocking.

The “hard wiring” of some silicon semiconductor devices can be reconfigured by software instructions, and this has led to some unexpected possibilities. An account in *Science* magazine put it this way:

The innovation arrived, so the story has it, as an epiphany. Hugo de Garis, a computer scientist who...describes his avocation as building brains, was visiting George Mason University in Fairfax, Virginia, in the summer of 1992, when he had a discussion with an electrical engineer. This “E.E. guy,” ... was telling him about computer chips known as field programmable gate arrays, or FPGAs, which are, in effect, pieces of hardware that can be reconfigured by software. “You can send in a software instruction, and it tells this programmable hardware how to wire itself up,” says de Garis.

De Garis then had a moment of visionary clarity: “The idea occurred to me,” he says, “that if you could send in a software instruction to a piece of hardware to wire it up, maybe you could look on that software instruction as the equivalent of a genetic chromosome. You could breed it, by mating chromosomes, mutate it randomly, maybe actually evolve the hardware. So then I started asking this guy, ‘Could you do this infinitely; could you just keep sending in instructions, rewriting the hardware again and again and again?’ And he said that for some FPGAs that was possible.” (Taubes, 1997)

The story goes on to describe a number of examples of this “silicon evolution”, emphasizing its engineering applications. But the most interesting example is not particularly exciting as engineering:

Perhaps the most intriguing variant of the evolvable hardware idea is one pursued by Adrian Thompson and his colleagues at the University of Sussex in the United Kingdom. Thompson thinks computer scientists are restricting the powers of evolution unnecessarily by putting it to work only on digital logic gates. “Silicon has a much richer dynamical behavior than just flicking between ones and zeros,” he says. Chip components can adopt a whole range of values intermediate between the standard 1 and 0...

The point here is that Thompson is making use of the fact that these FPGAs are able, in certain hard-to-find configurations, to program rules other than straight zero-one logic. He proposes to use the evolutionary technique to get a chip to find a configuration that will allow it to do a simple non-digital task…

So, rather than making their FPGAs follow the rules of digital design, says Thompson, “we don’t tell evolution anything about how we expect it to work. We let it find the best way.”

Thompson’s demonstration task was to feed an FPGA a single input—a 1-kilohertz or 10-kilohertz audio tone—and evolve it to generate a signal identifying the input: 1 volt for 1 kilohertz, 5 volts for 10 kilohertz. “This is actually quite a hard task,” says Thompson. “The problem is the logic gates are incredibly fast. They respond on a time scale of 2 or 3 nanoseconds. The input is orders of magnitude slower. So we were asking incredibly fast components to do something much slower and produce very nice, steady output,” he says. “All evolution had [for the task] was this little bit of silicon on a chip. ... It had to use the natural physical properties to get the job done.”

The system that resulted from the evolutionary process worked efficiently, says Thompson, but he didn’t have the “faintest idea how it worked.” [my italics] Back-engineering failed to decipher it, but what he did learn was that the circuit seemed to be relying on only 32 of the 100 available logic gates to achieve its task, and some of those working gates were not connected to
the rest by their normal wiring. “It was either electromagnetic coupling,” Thompson explains, “which is basically radio waves between components sitting right next to each other, or they were somehow interacting through the power-supply wiring.”

What has Thompson done here? He has taken a relatively simple device, a 100-gate FPGA, and through an evolutionary process caused it to learn a relatively simple task: to discriminate between two input frequencies. The machinery looks simple and the behavior corresponds to a very simple model. But the physical details – how the system is actually doing the job – are deeply obscure. This is not an isolated example. Workers on complex systems, artificial as well as natural, increasingly find that these systems may behave in orderly and understandable ways even though the details of how they do it are incomprehensible.

My question is: why should the brain be any different? It has also evolved, and continues to rewire itself during learning in a process that closely parallels evolution. So why on earth should we expect “back-engineering” – purely neurobiological investigation of the physical machinery – to yield a coherent account anytime soon? On the other hand, if the selection pressures are consistent (as in the FPGA-array example), we might well expect that the behavior of the system, its function, will in fact be understandable at the level of a simple model. So, I contend, if you want to understand behavior, begin by trying to understand its rules of operation.

Complexity. Complexity poses problems, by itself, of course. The most complicated objects currently made by human beings are computer chips. As I write, the Intel Corporation introduced a new chip with 3.4 million “gates” (transistors). By the time you read this, the latest chip will have more than 10 million gates. The connections among all these gates are of course perfectly well defined and each gate is (supposed to be) completely independent of all the others – there are no “global” factors, like circulating hormones, that allow events in one part of the chip to affect other parts in a graded way. The only effects are through well-defined wires. Even so, the task of “reverse-engineering” such a chip – figuring out from its input-output properties and from what can be seen of the internal wiring exactly how it works – is formidable. A few companies in fact make their living by doing just this, making “clones” of Intel microprocessors. Nevertheless, even though the fundamental principles of chip design are well understood, even though these companies can take apart an indefinite number of Intel chips, and even though they can be absolutely certain that every Intel 80486 they buy is identical to every other, the “clone makers” do not always succeed. Only very recently have clones of Apple's computers begun to appear, for example. Reverse-engineering is simple in principle, but often difficult in practice.

Reverse-Engineering the Brain. Reverse-engineering the brain is much more difficult than reverse-engineering even the most complicated computer chip, because (a) The brain has many more parts. The unit of brain function is the nerve cell or neuron. (At least, we think so—but because of long-range effects, there may be no real functional unit.) There are many differ-

---

45 Here’s another way to look at it. Suppose that Thompson instead of working in a university was part of the UK defense establishment. And suppose that his evolved device was stolen and smuggled to a foreign power with no information on its function. After months of scrutiny by microelectronics experts, what kind of answer could they give to their political interrogator when he asks “What does it do?” If they are good experimenters they will answer “It discriminates between two frequencies.” Probably, the interrogator will be satisfied at this point and it will only be a curious scientist who asks, “But how does it do it?” The simplest answer to give him, of course, will be “By evolving that way…” Neurobiology may have to settle for this answer for a long time…

46 See Staddon (in press) for a summary of previous work arguing for the “Darwinian metaphor,” the analogy between learning and evolution.

47 Much of the skill in designing these chips involves the elimination of unwanted interactions.

48 They cannot make exact copies, of course, for copyright reasons. What they hope to do is make chips that function in the same way as the Intel chip, without copying the circuitry exactly. That’s just what psychologists hope to do with the brain – we’re not trying to duplicate every neuron, just come up with models that duplicate the functions.
ent types of neurons, and there are also other types of cells in the brain, notably glial cells, that are thought to have subordinate functions\(^{49}\). There are about \(10^{11}\) neurons in the human brain (for comparison, there are about \(2.7\times10^7\) people in the United States and \(1.4\times10^6\) gates in the 80486 chip). The number of glial cells is even larger. (b) Neurons are complex structures that have very many connections with other neurons – the average is about 1000 connections according to a standard textbook on neuroscience (Kandel, Schwartz and Jessell, 1991). Each neuron is itself complicated: “The information carried by the neuron is encoded in electrical signals that travel along the axon and into the nerve terminal. At the synapse these signals are carried by one or more chemical messengers across the synaptic cleft. None of these chemical messengers carries unique information, like RNA or DNA. Indeed, some of them have several functions within cells as metabolites in other biochemical pathways. The corelease of several neuroactive substances from a presynaptic neuron and the concomitant presence of appropriate postsynaptic receptors permit an extraordinary combinatorial diversity of information transfer.” (J. H. Schwartz in Kandel, Schwartz & Jessell, 1991, p. 224.) (c) There are on the order of \(10^{14}\) synaptic connections in the human brain. Most researchers assume that all changes in behavior caused by experience – including learning – are traceable in some way to changes in synaptic connectivity. Some of these changes are permanent, others are more or less transient. But there is as yet no guarantee that learning depends on synaptic changes alone. (d) Interactions between individual neurons are probably not the whole story of brain function. In addition to local effects, there are also diffuse, long-range effects in the brain. Neurochemicals, such as some neurotransmitters and hormones, can affect many synapses at once. Graded electrical potentials may also have long-range effects. The nonlocal effects of neurotransmitters are just beginning to be explored. The role of long-range effects in overall brain function is still largely unknown. (e) Glial cells do not participate directly in the electrical signaling process carried out by neurons, but they probably play some role. The task of building up a picture of how behavior is generated by the brain from ever-deeper studies of local mechanisms – from channels to receptors to synapses to simple neural circuits to more complex circuits to the brain as a whole – will be monumental.

The fact of neural complexity implies that understanding the behavior of a whole organism by beginning with individual neural units (the so-called bottom up approach) is likely to be a slow process. Some progress along these lines has been made through the study of very simple organisms with limited numbers of neurons, such as the sea slug Aplysia and the nematode worm Caenorhabditis elegans. Considerable advances have also been made in studying particular parts, particularly peripheral parts, of the brains of higher animals: sensory systems such as vision and audition, for example, and some motor systems. In higher organisms, progress has been slowest in understanding processes that depend strongly on history, processes like learning where, by definition, behavior now may depend upon experiences long ago in an organism's past.

I believe that input-output modeling, the top-down approach, is a better strategy for understanding learned behavior. It may also be a necessary preliminary to understanding brain-behavior relations. Without cracking our Intel chip, we can learn a very great deal about how it works just by applying different input patterns and measuring the resulting output patterns. Once again, however, the clone maker has several advantages over the psychologist. Most importantly, he knows what to measure. In psychological terms, the electronic engineer begins with a pretty good idea of what constitutes stimulus and response: he knows that electrical impulses, particularly their spacing (not so much their amplitude) are all that is important. The psychologist is much less sure how to define his stimuli and responses. Is the “red light” the stimulus? Or light of a certain wavelength (color is not the same as wavelength)? Should the stimulus be

---

\(^{49}\) The functions of glial cells may be less subordinate than previously thought. See the 1994 Nature comment “Glia and neurons in dialogue” by David Attwell.
defined in physical terms at all? Psychologists and philosophers of mind have puzzled over these issues for many years.

The psychologist also faces special problems posed by the fact that we are dealing with biological material. Biology implies variability – not in every variable (for example, body temperature varies little in endotherms, skilled movements are carried out with great precision, etc.) and not always – but under very many conditions. There are good evolutionary reasons for variability, as we will see later on. But it means that brains, even the brains of identical twins, unlike computer chips are not identical. Even if the brains of twins were identical at some early point in their history, it is impossible to give two individuals – even individual rats, nematodes or sea slugs, much less human beings – identical life experience. Given different histories, even two initially identical brains will become slightly different – they will be “storing” different memories, after all. This difference, and what it implies, namely that experiments cannot always be replicated exactly, poses severe difficulties for a researcher who wants to understand how organisms are affected by different histories.

Repeatability. The chip engineer faces none of these problems. All 80486 chips are the same: chips are all identical replicas of one another, no problem with variability. And experience doesn’t change a chip irreversibly. Although every chip has a memory, the memory is volatile: that is, it requires electrical power to retain information. If the power is cut off, the memory is reset to its initial state and all the stored information is lost. So, any experiment with an individual chip can easily be repeated, either by resetting the memory of the chip and beginning again, or by taking another, identical, chip off the shelf. Repeatability is essential to the experimental method, and I will discuss two ways to ensure it in a moment. The point, for now, is that repeating an experiment on behavior is much trickier than repeating an experiment with dynamic-memory computer chips. The psychologist has only imperfect replicas (different organisms, even identical twins, are not really identical), and the effects of past experience cannot always be erased (animal memory, unlike computer memory, is not volatile).

How do experimenters deal with the problem of repeatability? There are two experimental methods, within-subject and between-subject. I devote the next section to a discussion of these methods at a fairly elementary level – even though most readers will already be familiar with them. Why take up space with boring methodological discussion? For two reasons, one general, the other specific to this book. The general reason is because methods tend to take on a life of their own, to be pursued not necessarily because they are the best approach but because they are the familiar, the accepted approach – and (in psychology) because they are suitable for the application of statistics. If experimental methods are not to be misused, it is essential to understand their limitations. The specific reason is my belief that the within-subject method is essential if our aim is to understand learning at the most fundamental level. But it must be accompanied, as in the past it has not been, by theoretical exploration.

**Within-Subject and Between-Subject Experiments**

The within-subject method is the method of physics, chemistry, physiology and some varieties of experimental psychology. The experimenter deals with a single piece of experimental

---

50 There may also be intrinsic variability, in the sense that even two identical individuals with identical experience, may nevertheless wind up with different physical brains. But the practical difficulties in studying brain-behavior relations are sufficiently immense even without this possibility.

51 The term “dynamic” — in a different sense than the one in this book — is also used.

52 The classic account of within-subject research in learning psychology is Murray Sidman’s 1960 book *Tactics of Scientific Research*, which is still well worth reading. Between-subject research, which is the dominant method in psychology, is covered in numerous statistical books. For a simple account see almost any introductory psychology textbook. For more recent account, see Staddon (2019).
material – a pendulum (say) for an early physicist, a couple of compounds for a chemist, a nerve fiber for a physiologist, a human being or a pigeon, for a psychologist. Experiment involves repeated manipulations on the same subject matter: the pendulum is pushed, compounds are heated or mixed, the neuron is stimulated, the pigeon or human being is subjected to a series of training and testing procedures. Let’s look first at a simple physics experiment, then at a psychology experiment, to expose the common features.

The pendulum experiment

First, the most difficult part: framing an experimental question that is both answerable and interesting. Finding a question that can be answered experimentally is much easier than finding an answerable question that is also interesting, i.e., whose answer can lead to new developments in the field – not just dotting a few i’s or crossing a few t’s. We need to decide first what things we should measure and then what manipulations we should make. A pendulum has a length, a weight, and a distribution of weight (is it just a stick, or is most of the weight at the end, in the bob?). It also has a color, a material composition, a hinge or string, a height above the ground, and numerous other properties that are not obviously related to its behavior, which is pretty simple: it swings back and forth with a certain period (time between swings) and it persists in swinging for a longer or a shorter time after an initial push (decay time). The thing we want to measure, the dependent variable, will be something to do with its swinging. The most comprehensive measure would be the entire pattern of movement in time, beginning with an initial push. But if decay time can be neglected, we can just measure the period, precisely defined as the time it takes to return to the same position after an initial displacement.

The next question is: “what should we do to the pendulum?” – what manipulations should we make? Pendulums fortunately have pretty simple behavior (that’s why physics is farther ahead than psychology): about all you can do with a pendulum is push it or (which amounts to the same thing) displace from the vertical by a certain angle. The angle of initial displacement is the independent variable, the thing which is manipulated by the experimenter.

Once we have decided on what to measure and what to do, we can ask the critical question: is the within-subject method indeed an appropriate way to study this system, as we have defined it (i.e., in terms of these dependent and independent variables)? The answer has to do with repeatability and a property called history-independence53: Does the pendulum behave the same way beginning with the same initial conditions? For example, suppose we always begin our experiments with the pendulum at rest. If we give it the same push (e.g., displace the bob through the same angle \(\delta\)), does it always follow the same pattern of movement? If it does, if the behavior is repeatable, then we can ask some experimental questions. How does the decay time depend on the material of the string? Of the bob? How does period depend on the weight of the bob? On the length of the string? On the size of the initial displacement?

---

53 The term path independence is also used.
All experimental questions are not equal. Of these questions, only the last three turn out to be interesting—in the sense that they reveal the working of general physical laws. So let’s manipulate three things: the bob weight, initial displacement and the string length and see how they affect period. To study the effect of length, we begin with a given bob weight, say 100 gm, and vary string length. At each string length, we displace the bob by the same fixed angle, $\delta$, let it swing, and measure how long it takes for the bob to return to the starting point (i.e., the period). We do this for several lengths. The result of such an experiment will be a table of numbers showing how period depends on string length. The numbers from one such hypothetical experiment are graphed in Figure 1.2 (the units on both axes are arbitrary).

The data are represented by the filled squares. They seem to follow a smooth curve, and I have fitted a simple power function to the points. It that follows the equation: $T = k\sqrt{l}$, where $T$ is the period, $l$ is string length and $k$ is a constant. There are in fact several other negatively accelerated (concave downward) functions that would fit these data almost equally well. But most of them are more complicated, and the square-root function is the correct one, because it alone can be deduced from the basic principles of Newtonian physics.

What about bob weight and initial displacement? The same experimental design will do for these also: settle on a string length, then try different bob weights and displacements and look at the period. The result here is much simpler, though: bob weight has essentially no effect on period and neither does displacement, providing it is small. All these results, the independence of period and bob weight, the equivalence of all small displacements, and the square-root relation between period and string length, were deduced by Isaac Newton from his laws of motion.

What can we learn from this trivial example? (a) The within-subject method is appropriate when observations are repeatable: the same experimental conditions (string length, bob weight, displacement) always produce the same result (period). (b) Repeatability implies history-independence: we always get the same input-output relation (length vs. period relation), no matter what lengths we have tried and no matter what, or how many, observations we have made. (c) Any set of data can be fitted by several (actually, an infinite number of) mathematical functions. The decision among them must be made initially just on grounds of simplicity (parsimony) and the degree to which each fits the data. But the final decision will depend on theory: the best function is the one that both fits well and can be deduced from some more general theory that relates this result to others—as Newtonian physics relates the behavior of a pendulum to the motion of the planets and the trajectories of falling bodies.

This example illustrates the usual sequence in basic science, which involves several steps, almost none of which are algorithmic, that is, reducible to a well-defined set of rules. The first step is usually inductive—find some empirical regularity. There are few rules for this. Charles Darwin believed that he set out “in true Baconian fashion” and let the facts speak for themselves. But as many philosophers have pointed out, fact seeking must usually be guided by some explicit or implicit notion of what is important—a tacit theory, at least. (On the other hand, the first in-

---

54 I will use the terms theory and model more or less interchangeably. Theory has somewhat grander implications than model: a model is assumed to be both less ambitious and less well established than a theory. We speak of Newton’s theory (of gravitation) but Estes’ model (of spontaneous recovery), for example. But conceptually, the two are the same.
ductive step in the pendulum example was the discovery of the principle of the pendulum, which is lost in prehistory and owes little to any pre-existing theory, either tacit or explicit.) The inductive step in the pendulum example was to plot period vs. string length. The next step is to see if there is some quantitative way to summarize this regularity. For the pendulum example, I did some curve fitting which, these days, *is* often algorithmic. There are numerous computer programs which will find the best fitting function from a set of possibilities, for example. But this step still leaves you pretty close to the data. The next step is the critical one, often omitted in psychology. It is to see if the regularity you have found can be generalized in some way. Is the square-root function for the pendulum an example of some more general principle? Charles Darwin (1809-1882) famously said “The line of argument often pursued throughout my theory is to establish a point as a probability by induction and to apply it as hypotheses to other points and see whether it will solve them.” By “point” here, Darwin of course meant not the raw data, but a generalization based on the data, not “this bug has six legs” but the generalization “all insects have six legs.”

The journey from empirical regularity (the square-root law) to (successful) theoretical generalization (Newton’s laws of motion) is a difficult and mysterious one that is made relatively infrequently in science. Millions had seen the variety of nature, the perfections (and imperfections) of adaptation and the prodigality of reproduction before Darwin connected them through the principle of natural selection. The potential generality of Gregor Mendel’s model for his data from breeding sweet-peas was not recognized for more than thirty years, when “Mendelian genetics” was re-discovered by three other groups. In a criticism of Francis Bacon (1561-1626), who overemphasized the inductive step in the scientific sequence, philosopher Bertrand Russell (1872-1970) commented:

Bacon’s inductive method is faulty through insufficient emphasis on hypothesis. He hoped that mere orderly arrangement of data would make the right hypothesis obvious, but this is seldom the case. As a rule, the framing of hypotheses is the most difficult part of scientific work, and the part where great ability is indispensable. So far, no method has been found which would make it possible to invent hypotheses by rule. Usually some hypothesis is a necessary preliminary to the collection of facts, since the selection of facts demands some way of determining relevance. Without something of this kind, the mere multiplicity of facts is baffling. (1946, 556-7)

Russell was writing at a time when the research enterprise was a tiny fraction of its present size. His comment is even more true today than it was fifty years ago. Anyone who has tried to master the scientific literature in a growing area like neuroscience will emphatically agree with Russell that “the mere multiplicity of facts is baffling.” Because U.S. science, particularly, has followed a largely Baconian road, facts tend to multiply unchecked. This causes little trouble in fields like physics and molecular biology, where the theoretical underpinnings are secure and constraining. But in psychobiology the bias for experimental facts over integrative theory has led to a tantalizing morass: thousands of meticulous and often ingenious experiments whose net contribution to human understanding is often considerably less than the sum of its parts.

**The Reflex Experiment**

Now for a psychology experiment. If you shine a bright light in someone's eye, the pupil constricts. This is known as the pupillary reflex. It helps control the intensity of light falling on the retina, just as the diaphragm of a camera controls the amount of light falling on the film. The

---

56 To be fair to Bacon, who was a brilliant man, albeit a cruel and deceitful one, he was well aware of the need for hypothesis, but had few examples to point to, given the relatively primitive state of science in his day.
The pupillary reflex is an example of a *stimulus-response* relationship. The independent variable (the stimulus) is the intensity, wavelength and duration of the stimulating light; the dependent variables (response properties) are the *latency* (i.e., time after the stimulus onset when some response can be detected), *duration* and *magnitude* of the pupillary constriction.

If the reflex is elicited over and over again, the response will diminish — from fatigue or habituation. But after allowing a time for recovery, it will return to its original level. With this constraint, the pupillary reflex seems to be repeatable, just like the behavior of the pendulum. Hence, we can study it in a single subject. We might look at other subjects just to make sure that our subject is typical, but the properties of the reflex can be completely worked out with only one. Like the pendulum experiment, the data from this one can be summarized graphically, by a plot of response magnitude vs. stimulus intensity, for example.

**Irreversibility, Hysteresis and Associative Learning.**

Now let's look at both these experiments again and note a couple of things that might make the within-subject method less useful. Every child learns that one way to break a piece of wire is to bend it back and forth repeatedly. At first, there is considerable resistance, but after a while resistance decreases and eventually the wire breaks. The effect is termed metal fatigue, and it reflects motion-induced changes in the crystalline composition of the metal.

**Irreversibility.** Suppose that our pendulum bob hangs not from a piece of string but from a wire. As the pendulum swings, the wire is bent. What will be the effect of metal fatigue on the results of our experiment? It will show up first as a failure to replicate results. Because the wire is stiffer at the beginning of the experiment than at the end, after many bendings, the period early on will be shorter than later. Because the stiffness of the wire decreases with “experience” the pendulum will swing more freely later in the experiment. Consequently, a repeat of an experiment with the same wire will give a different period. Hence, an exact description of the results will have to include not just the current pendulum length but also information about the *history* of the system: how many swings has this particular wire undergone? How big were the excursions (the *amplitude* of the oscillations)? Over what period? And bob weight will no longer be irrelevant, because the heavier the bob, the more the wire will be bent, because the pendulum will go through more swings after a given displacement: the response will have a lower decay rate. Metal fatigue is an example of an *irreversible* effect: after *N* excursions, the wire will be in a unique *state* different from its *initial state* — and the initial state cannot be recovered.

**The Between-Subject Method.** How can we study irreversible behavior like this? The between-subject method deals with the problem by using more than one experimental “subject”. To study the effect of wire length on period, we will need several pieces of wire of the same “experience” (i.e., with the same fatigue history). Given these replicas, we can compare the periods associated with different lengths knowing that the irrelevant historical factor does not contaminate the results. We have “controlled for” the effects of history.

Effects like fatigue are often quite *variable*, that is, two apparently identical pieces of wire with the same history of bending may nevertheless have different stiffnesses. If we cannot measure stiffness directly, the only way to do the pendulum experiment is to use more than one piece of wire in each experimental condition. Thus, we might have 100 pieces of 20-cm wire, 100 40-cm pieces and so on, and each set of 100 pieces will in turn be divided into 10 groups of 10, each with a different history. (Each of these groups of 10 is called a *sample*.) The data would then have to be averages of the period under each condition (i.e., each combination of

---

57 Of the camera designer or his biological counterpart, natural selection.
length and history). Given a big enough sample, the group average will be reliable in the special sense that if we were to repeat the experiment with another sample of ten, we can expect to get about the same average – even though the individual pieces in each sample will all be different. This gain in reliability is a consequence of a fundamental statistical property known as the central limit theorem: If you generate numbers at random, the average of a sample of N numbers will (a) vary much less than the numbers themselves and (b) tend to a normal (bell-curve) distribution, no matter what the distribution of numbers themselves.

Using the within-subject method with a historical system. The irreversible changes caused by bending make the wire pendulum a historical system, that is, a system whose behavior depends on its past experience. It has long been a convention in psychology that historical systems must be studied with between-group methods, as I just described. There is a cost to the between-group method, however. More subjects means more work, but the main problem is that the behavior of a group average need not reflect accurately the behavior of any individual in the group. The average American family has 2.4 children, but we would be alarmed to find a fractional child in any family. The average performance of a group of subjects learning some task always improves smoothly, but individual subjects may go from zero to perfect in one trial (see Box 1.1).

The Theoretical Approach. Fortunately, there is an alternative approach to the study of historical systems that preserves the advantages of the within-subject method. I’ll call it the theoretical approach because it uses theoretical exploration to discover the causes of unrepeatability. The metal-fatigue problem provides a good illustration. Recall that the problem showed up first as a failure to replicate results: given the same pendulum and the same displacement, the measured period was different on the two occasions. What to do? We can simply accept the fact of unrepeatability and go over to the between-group method – or we can try and understand the reasons for the failure to replicate. Unfortunately, the theoretical method gives no specific pre-

---

**Box 1.1: Where is the variability?**

The between-subject method is often thought to be theory-free, but the value of the results so obtained depends very much on what the underlying process really is and what are the sources of error (usually termed noise). For example, suppose that the real relation between \( x \) (the independent variable) and \( y \) (the dependent variable) is a power function,

\[
y = Ax^B,
\]

where \( A \) and \( B \) are parameters. Suppose that all the error in measurements of the dependent variable, \( y \), comes from variation in parameter \( A \). Will an arithmetic average value for \( y \), when plotted against \( x \), give an accurate picture of the underlying power function? A little thought shows that it will. For example, suppose that we make just two measurements of \( y \) (same \( x \)), and the two values for \( A \) on those occasions are \( A_1 \) and \( A_2 \), so that

\[
y_1 = A_1x^B \quad \text{and} \quad y_2 = A_2x^B.
\]

The average value for \( y \) is thus

\[
\frac{1}{2}(y_1 + y_2) = \frac{1}{2}(A_1+A_2)x^B,
\]

which is still a power function with the correct exponent, \( B \), and a multiplier at the average of the actual values. Thus, in this case, the averaging process does not distort the true form of the underlying theoretical relation between \( x \) and \( y \). The same is true if the noise is simply additive, i.e.,

\[
y = Ax^B + \varepsilon,
\]

where \( \varepsilon \) is a random variable with mean zero: the relation between \( x \) and the average value for \( y \) will approach the true relation as sample size increases.

On the other hand, suppose all the error is in \( B \)? In this case, the measured relation will be

\[
.5(y_1 + y_2) = .5Ax^B + x^{B_1} + x^{B_2}
\]

which is not a simple power function. The problem can be solved for the power function by using the geometric mean, \( \sqrt[y_1y_2]{y_1y_2} \), rather than the arithmetic mean, but the solution is not so simple for other theoretical relations or if there are several sources of variability. And this method is of no help if the theoretical function is unknown.
scription for what to do; it is not algorithmic. One reason the between-group method is so popular is that it guarantees a result; numerous texts describe exactly how many subjects are needed and what statistical tests are to be used for every possible occasion, and computer programs provide painless calculation. If you follow the prescription, reliable (but not necessarily useful or meaningful) results are guaranteed. The theoretical method provides no such security. What it requires is simply a bright idea: knowing the problem area, knowing the properties of the material you're dealing with, it is up to you, the scientist to come up with a hypothesis about the causes of unrepeatability. The next step is to test that hypothesis directly. If your guess is wrong, you must guess again.

The Theoretical Axiom. The assumption behind the theoretical method is that a failure to get the same result when all measurable conditions are the same, implies the existence of a hidden variable (or variables). We may think that conditions are the same on the second experiment as the first, but if the period we measure is not the same, then something we are not measuring must have changed. What that “something” is in the pendulum experiment is pretty obvious. It is the stiffness of the wire. The data suggest that stiffness changes with experience and affects period. If you can figure out a way to measure stiffness, you can then go on to see how period depends on it. If the relation is one-to-one, a given stiffness (and wire length) corresponds to a given period. Experimental repeatability is restored: given a certain pendulum length and a certain stiffness of the wire, the period is fixed. The system has been converted from a historical to an ahistorical one. And an ahistorical system can be studied using within-subject rather than between-subject methods, which are desirable for the reasons just given. (It is worth remembering that almost all the discoveries of physics have been made using the within-subject method and without the aid of statistics.)

Parameters and Variables. The wire pendulum system is characterized by one parameter (wire length) and three variables: stiffness, and the position and velocity of the bob. A parameter is just a variable that changes slowly enough to be neglected during an experiment – stiffness might be considered a parameter rather than a variable if the experiment is short or the wire is of a type that fatigues slowly. Variables are the things that define the state of a system. That is, if we know the values of the variables, future behavior of a (deterministic) system is perfectly determined. These terms will mean more after I have discussed more examples.
Notice the special advantage of the approach through theory: by trying to understand the causes of unrepeatability rather than just accepting it, something new has been learned about the nature of the system. The between-group method deals with variability and unrepeatability by brute force. And all it guarantees is that you have found out something about an “average subject.” What you find may nevertheless not be true of any individual. The within-group method deals with unrepeatability by seeking its cause, which usually means changing the experimental question. To be sure, there are many cases, particularly in applied areas, where we are not at liberty to switch questions. In these cases, the between-group method is the only option. For example, given an incurable disease and two marginal treatments for it, there is no short-term alternative to trying the treatments, plus a placebo “treatment,” on three large groups of patients, in the hope of finding the better alternative. But we learn little about the disease from such an approach. Consequently, in basic research, where we have the leisure to choose our own questions, the theoretical approach is essential.

**Hysteresis.** Metal fatigue is an irreversible change. Under some conditions, the pupillary reflex shows a less drastic kind of history-dependence that allows for a mixed experimental strategy. Figure 1.3 shows the results of three hypothetical experiments. In all experiments, a brief “pulse” of light is presented to the eye (inset) and the response magnitude is measured (inset arrows). In a series of trials, the intensity of the pulse is progressively increased and then decreased: 10 increasing stimuli and then the same 10 in reverse (decreasing) order. The single line at the top is the experiment just described, in which the same stimulus always gives the same response. Trials (stimulus presentations) in this experiment were very far apart. Now look at the pair of light lines with open squares. This is the same procedure, but with trials a bit closer together. Notice the difference between the ascending and descending sequences. In the ascending sequence (upper line), as stimulus intensity increases, the response also increases, but at a slower and slower rate. But in the descending sequence (lower line), as stimulus intensity decreases, the response drops rapidly at first but then more slowly. The result is that a given stimulus intensity does not always produce the same response – a failure to replicate. The kind of unrepeatability represented by closed graphs like this is called hysteresis.

The pair of heavy lines with filled squares is another repeat of the same experiment, but with an even shorter time between stimuli. Notice that the two curves are now farther apart and both turn down before the end of the stimulus series – a paradoxical result, because a stronger stimulus then gives a smaller response.

What might be the reason for the hysteresis here? Recall that many reflexes show reduced strength after they have been excited. I use the term reflex strength in a special sense here, to refer to the relation between stimulus and response (I'll give a more exact definition in Chapter 5). For example, suppose we present a stimulus of a certain intensity and get a response of magnitude 100 (the units don't matter). Suppose we present the stimulus again, immediately, and the response now is only 50. There are several possible reasons (fatigue, habituation, adaptation), for post-response reduction in strength. I will get to them in later chapters. But for the

---

58 Much nonsense is written about the supposed lack of difference between applied and basic research. The distinction is this: in basic research the main aim is to find the right questions. In applied research, the questions are given. The methods are similar, but the aims are very different.

59 Note that if the treatments were not marginal, if one or other were close to 100% effective, there would be no need for statistical comparisons. Penicillin, which was originally highly effective against a wide range of bacterial infections, was universally accepted with no statistical comparisons whatever, for example. Elaborate statistical designs, and large numbers of subjects, are needed only when a small effect must be distinguished from no effect at all. The method has run into problems that should, perhaps, have been obvious at the outset (Ioannidis, 2005; Meehl, 1978; Staddon, 2017).

60 The term hysteresis was first used in connection with the study of magnetism, in physics (in magnetic hysteresis, the up- and down-series in Figure 1.6 are reversed, however). The word is taken from the Greek hysteron, meaning womb, because of the oval shape of the closed graph. The word hysteria comes from the same root.
moment, just notice that this reduction in strength often depends on interstimulus interval. The reflex has to “recover” after being stimulated. Consequently, if we wait a long time in between stimuli, the reflex has fully recovered, and the same stimulus always gives the same response. But if we wait a shorter time, it has not recovered completely, and the strength is reduced.

It is this reduction in strength that produced the hysteresis in Figure 1.3. At the end of the ascending sequence, the reflex is “fatigued” because it has been stimulated often, most recently with intense stimuli. Thus, when the descending sequence begins, reflex strength is reduced and the same stimulus elicits a weaker response, yielding the oval “hysteretic” pattern. The shorter the time between stimuli, the larger the effect and the fatter the graph. Thus, this complicated-looking graph is just a consequence of a simple effect: a transient reduction in strength of the reflex every time it is excited.

Hysteresis doesn’t rule out the within-subject experimental method, but it does mean that we need to wait a while between trials to be sure that the reflex has been restored to its initial state. Many simple types of adaptive behavior retain the effects of a stimulus for a brief time in this way. The term short-term memory (STM) is often used to refer to this property.61

**Associative Learning.** Today, a pupillary-response experiment would be carried out entirely automatically. A computer would turn the stimulus on and off at precisely programmed intervals. This is an improvement over the old, manual method because it gives the subject no warning. In earlier days, however, the stimulus would have to be turned on by the experimenter, who might actually be in the same room with the subject. The experimenter’s activity could therefore provide the subject with a reliable warning that the stimulus is about to be turned on. This relation between an irrelevant stimulus and something of importance to the organism may remind you of a more famous experiment from years past. Russian physiologist Ivan Pavlov encountered the unexpected results of just such a pairing between a “neutral” stimulus and a reflex stimulus in his experiments on the digestive function of dogs nearly ninety years ago. In Pavlov’s case the neutral stimulus (now termed a conditioned stimulus or CS) was just the activity of placing the experimental animal in a restraining harness. The reflex was salivation (termed the unconditioned response, UR), elicited by food powder (the unconditioned stimulus, UCS) placed in the animal’s mouth. Pavlov noticed that his dogs very soon began to salivate in advance of the food, just in response to the preparations. He tested his conjecture that the pairing between the preparations and the food was responsible for this new learned “reflex” by explicitly pairing different neutral stimuli – a bell, the click of a metronome, and many others – with the delivery of food. He confirmed that the reliable CS-US sequence invariably caused a response (the conditioned response) to the CS that usually resembled the unconditioned response. This phenomenon is now known as classical, Pavlovian or respondent conditioning.62 The connection established by the Pavlovian procedure between the conditioned stimulus and the conditioned response is an example of associative learning.

In the pupillary reflex, if the unconditioned stimulus (the light pulse) is reliably signaled by some neutral stimulus (like setting up the stimulating apparatus, or sounding a tone), after a few pairings the neutral stimulus will begin to elicit the response on its own. This acquired association between stimulus and response is associative learning, a type of change more persistent than hysteresis and different in other respects as well. Associative learning falls somewhere in between hysteresis and irreversibility in terms of its persistence. There is no doubt that some associations leave no lasting trace whatever: a phone number learned and used once is soon forgotten, for example. On the other hand, many childhood experiences are never forgotten. In

---

61 The term STM is used in a number of other ways as well – to refer to human’s ability to recall recent events, for example, and as the name for mathematical processes that compute different kinds of running average (e.g., the leaky integrator discussed at length in later chapters).

62 These three terms are essentially equivalent. Needless duplication of terms – a sort of academic niche differentiation (or product differentiation, as an advertising expert might put it) – is a chronic disease of scholarship.
conditioning, even though the conditioned response can easily be eliminated through experimental extinction – repeatedly presenting the CS alone without the US – the subject is certainly not in the same state afterwards as he was before. The conditioned response can be reacquired much more rapidly than it was learned the first time, for example. Associative learning often has irreversible effects and is usually studied by between-group methods. The approach through theory can also work and I describe some examples in later chapters.

**SUMMARY**

Complex systems that have evolved via natural or artificial selection may behave in a simple way, even though the physical details of their operation are obscure. It is likely that the neurophysiological basis for history-dependent behavior is extremely complex and may never be completely unraveled. Consequently, neurophysiology is unlikely to provide a shortcut to understanding learned behavior. History-dependence implies the existence of hidden variables. Since the prospects for uncovering these through neurophysiology are not good; and since the methods of introspection are philosophically suspect and are in any case inapplicable to animals, the most promising approach is through theoretical exploration — the invention and testing of parsimonious black-box models. Within-subject experiments are best for this purpose. These ideas summarize the approach of theoretical behaviorism.

---

**Box 1.2. Historical Note: Behavior Analysis and the Within-Subject Method**

In the psychology of learning, the within-subject method is associated with the behavior-analysis movement begun by B. F. Skinner (1904–1990) and his students and associates in the 1950s (see, for example, Ferster & Skinner, 1957; Sidman, 1960). Skinner discovered reinforcement schedules, procedures for arranging food reinforcement or electric shock punishment to be delivered contingent (dependent) on various properties of an animal’s behavior: number or time of responses, rate of responding, etc. A striking feature of reinforcement schedules is that the pattern of behavior that animals develop after prolonged exposure — so-called steady-state behavior — is more or less independent of the animal’s previous history. Reinforcement schedules can therefore be studied using the within-subject method. Nevertheless, simple tests show that an animal exposed to schedule B is not in the same state after the first exposure as after an intervening exposure to schedule A: the sequence BB does not leave the animal in the same state as the sequence AB, even though its behavior may look identical in both “B” conditions. The difference shows up (for example) if we add a third condition, C, which might be extinction (i.e., no reinforcement). By comparing different groups, it is possible to show that the animal will usually persist much longer in extinction in the B-behavior following the BB sequence than following the AB sequence. In extinction following AB, the A-behavior will after a while displace the B-behavior. Thus, the similarity of overt behavior in condition B in the AB and BB animals is misleading. They may look the same, but they aren’t really the same, because of their different histories.

Once facts like this became widely known, the behavior-analysis movement had only two real choices: either go to the between-subject method, or retain the within-subject method and look for hidden theoretical variables to account for historical effects. The first method has been adopted by a large school of animal-learning researchers, primarily those interested in classical conditioning. But the behavior analysts stuck with the within-subject method. Unfortunately, they also bought Skinner’s curious view of theory, which caricatured its role in science. Laurence Smith summarizes Skinner’s view as follows: “Skinner admits that ‘theories are fun,’ but he insists that the activity of conjecturing is less efficient than that of observing. Conjecturing appeals to the investigator ‘whose curiosity about nature is not equal to his interest in the accuracy of his guesses.’ Moreover, conjecturing is said to be ‘wasteful,’ to create ‘a false sense of security,’ and to lead to ‘useless’ and ‘misdirected’ experimentation, ‘bootless theorizing,’ and the necessary loss of ‘much energy and skill.’” (Smith (1986), p. 272. Some of this is true, some false but all is misleading. For example, to oppose curiosity and conjecture is to oppose the end to the means. Curiosity prompts conjecture and conjecture leads to experiments that may satisfy curiosity. Skinner might as well have written “makeup appeals to women who value their appearance less than their experiments with cosmetics.”
Chapter 2

ADAPTIVE FUNCTION, I: THE ALLOCATION OF BEHAVIOR

If one way be better than another, that, you may be sure, is nature's way. Aristotle, Nicomachean Ethics.

Box 2.1: Historical Note: Charles Darwin (1809-1882)

Charles Darwin was born to an affluent English physician and his wife on the same day as Abraham Lincoln. He grew up conventionally. He much enjoyed the outdoor life, to the point that in his late teens his father thought him “an idle sporting man” unlikely to achieve much. He attended medical school in Edinburgh, where he “put forth no very strenuous effort” to learn medicine. Later, destined for the clergy, he went to Cambridge University where he spent much time in hunting and riding – and, a foretaste of his future, collecting beetles and other natural curiosities – as well as talking with the botanist J. S. Henslow. He pursued his collecting to the point that his report of at least one new species was published in a pictorial collection. He said that he learned little at university. But he obviously learned something from Henslow, and from Adam Sedgwick, the geologist with whom he went on field trips. Like the rest of his family, he was devout as a youth, but, to his wife's sorrow, later lost any belief in conventional religion.

The formative experience of Darwin's life was his five-year trip on the Royal Navy survey vessel HMS Beagle, as unpaid naturalist and companion to its captain, moody young aristocrat James FitzRoy. Like many great scientists, Darwin's true education was directly from nature. Unlike Bacon’s typical scholar, he studied matter, not words. During landfalls – in Tenerife, South America, the islands of the Pacific – he collected rocks, plants, animals and birds in abundance. In the long sea-spells in between, he classified, wrote and thought deeply on the significance of his accumulating experience. In the Pacific he tested his first elegant theory, conjectured in England before the trip, on the formation of coral reefs. The voyage developed in him the “habit of energetic industry and of concentrated attention.” Like Sir Isaac Newton, when asked how he made his great discovery, Darwin could also have answered “by always thinking unto it.”

A few years after his return to England, well-supported by his father's money, Darwin retired to Down House (now a museum) in Kent – and became a semi-invalid, with an illness still not identified. He used his illness well: to avoid professional obligations and to organize his large family as assistants in his work. He wrote, corresponded and did numerous experiments – on climbing plants, pigeon breeding, on the survival properties of seeds in sea water, and many others. He worked for eight years on the classification of barnacles (his is still the standard work), to the point that his young son, visiting a friend's house for the first time, asked “And where does your father do his barnacles?”

Darwin was hesitant to publish on evolution for fear of religious censure and became almost obsessive in gathering evidence to buttress what now seems an already overwhelming case first for evolution and second, and most importantly, for natural selection as its chief agent. His reluctance to publish got a familiar comeuppance when, in 1858 he received from the young biologist Alfred Russel Wallace (1823-1913), recovering from fever in what is now Indonesia, a manuscript that contained an independent discovery of natural selection – even the name was the same – and its role in evolution. The story of how Wallace and Darwin amicably reported their work together (in an 1858 paper to the Linnaean Society, whose unremembered President at year’s end commented that “little of note” had occurred) has become a famous example of honor between Victorian men of science. (Some recent writers, ever-vigilant to discover historical victims of the class struggle, have argued that Wallace was treated unfairly by Darwin, but Wallace didn’t think so. He remained close to Darwin until Darwin’s death and when he wrote a book on natural selection he entitled it... Darwinism.)

In the year after their joint paper, Darwin hurriedly put together an “abstract” of the much longer book he had planned to write. This abstract was the 502-page Origin of Species, which was to influence scientists and scholars in fields as diverse as geology, biology, psychology and political science. (Karl Marx dedicated Das Kapital to Darwin and sent him a copy, which is still at Down House. The pages are uncut.) Darwin died in 1882 and is buried near Isaac Newton in Westminster Abbey.

In addition to the Origin and his lively account of the Beagle voyage, Darwin's short autobiography, written for his children, is a wonderfully direct account of his life and thoughts. Many insights can still be gained from his monumental The Variation of Animals and Plants under Domestication, which contains most of what Darwin had to omit from the Origin. Michael Ghiselin's The Triumph of the Darwinian Method is an excellent exploration of the subtlety, modernity and power of Darwin's theoretical thinking.

The form and behavior of animals often make sense as efficient ways of dealing with environmental problems. Pelagic fish are hydrodynamically efficient, for fast, low-energy movement in water; vultures have keen vision to spot their prey at long distance; all animals prefer a rich...
source of food to a poor one. All these characteristics are obviously adaptive, in the sense that they make it easier for animals to do things that favor their growth and survival. Other characteristics, from the peacock’s tail to the human appendix, pose more of a puzzle. It was Darwin’s genius (see Box 2.1) to see the common principles that link adaptive as well as apparently maladaptive features: natural (which includes sexual) selection and heredity, including the principles of ontogeny. Selection generally favors adaptation, but not always; and the principles of heredity place limits on what selection can accomplish. In other words, biological systems are not always adaptive, but they are adaptive often enough – “behavior reeks of purpose,” in Tolman’s memorable phrase – that adaptation has always commanded attention in its own right.

Most economists, and some behavioral and evolutionary psychologists and biologists, some of the time, try and understand behavior in terms of its adaptive function, apart from the process that gives rise to it. They do it by making optimality models. For economists, these are models of “rational” man; for biologists, they are models of perfect adaptation. For example, if you think that tea and coffee are partial substitutes for one another in the preferences of most people, you will predict that if the price of tea rises, not only will people buy less tea, they will also buy more coffee. Implicit or explicit constraints on perfection are implied. The analysis takes no account of other ways to spend the money, for example. No limits are placed on our ability to tell one beverage from another, although they must exist. Every optimality analysis has some limits, although these are often not explicitly acknowledged or recognized.

If you think that the value — utility — of most goods follows a negatively accelerated function, so that the more you have of something the less you value a bit more, then you will predict that people and animals will tend to show partial rather than exclusive preferences, allocating their resources so that they have some of each good rather than devoting everything to one good alone. Effects like this can be seen as rational – optimal – adjustments that maximize utility. Once the ideas of substitutability and utility are made explicit, straightforward mathematics allows predictions about how rational people should behave when costs or income change. This chapter and the next discuss the strengths and weaknesses of the optimality approach to adaptive behavior. The focus of this chapter is psychology, of the next, economics.

THE OPTIMAL ALLOCATION OF BEHAVIOR

Psychologists have tended to avoid dealing with notions like “purpose” and rationality” preferring causal accounts that explain behavior in terms of antecedent rather than consequent events. The ubiquitous behavioristic notion of reinforcement is an intermediate case that I will take up in a moment. But in recent years, economists have become more experimental and behavioral psychologists have acquired an interest in economics, and the result has been the growth of the field of behavioral economics (see, for example, Hogarth, 1990; Lea, Webley & Young, 1992; Rachlin, 1995).

Like most things in behavioral psychology, the application of economic ideas began with speculation about the nature of reinforcement. A seminal notion was David Premack’s (1965). He suggested two things: First, that an activity such as running, or playing – or even eating – is reinforcing in itself, quite apart from any stimuli (such as food) that may be involved. Second, he suggested that the reinforcing value of an activity is directly related to its frequency of occurrence: an activity that occurs 80% of the time is ipso facto more reinforcing than one that occurs only 10% of the time. Both these provocative ideas are more or less wrong. Heat, for example, is reinforcing to a cold animal quite independent of whatever activities may accompany its delivery; electric shock is punishing whether or not the animal cowers or freezes. And frequency is not the same as reinforcing value. Sexual congress may be highly reinforcing, though engaged

---

63 Pedants will point out that people are animals. I know; but it’s less longwinded to label “infrahuman” animals as just…animals.
in infrequently. Nevertheless, Premack’s novel approach prompted people to think more broadly about reinforcement and what it might mean.

In a typical experiment, Premack studied the reinforcing relations among the activities of Cebus monkeys. The monkeys were in individual cages and there were four things that they could play with: a lever (L), a plunger (P), a hinged flap (F), and a horizontally operated lever (H). Access to each of these things was controlled by the experimenter. Premack compared the proportions of total time spent in each activity under free conditions with the proportions after he had imposed a 1:1 contingency (10-s of Activity A bought 10-s access to Activity B) between the several possible pairs of activities. The contingency was like a fixed-ratio 1 (FR 1) schedule of reinforcement, but the units were time rather than responses. Activity A here is termed the ope

The clearest predictions possible were those for Chicko, who in the first procedure [free access to all activities] showed three reliably different response probabilities [proportions of time spent]. Indeed, Chicko's protocol made possible three kinds of contingencies: contingent response [the reinforcer] higher than, less than, and, in one case, about equal to the free [operant] response....the outcomes for the three types of contingencies were as follows: (1) contingent response higher than free response produced...an increment in the free response; (2) contingent less probable than free response produced...a decrement...; (3) the one case in which the responses were about equal produced little or no change, increment or decrement. (Premack, 1965)

Thus, a monkey that spent 10% of its time playing with L, and 20% playing with P under free conditions, increased its level of L when L, the operant response, had to occur for several seconds in order for P, the contingent response, to be available for the same number of seconds. Evidently, access to a high-probability activity will serve to “reinforce” a low-probability activity. The reinforcing effect of a high-probability activity on a low-probability one has come to be called the Premack Principle of reinforcement.

There is a more general version of the Premack Principle that takes account of schedules other than one-to-one. For example, suppose that response A takes up twice the time of response B, under free conditions. Response B may nevertheless reinforce response A if the schedule is greater than two-to-one: say 30-s access to A to get 10-s access to B. More formally, the response-deprivation hypothesis (Timberlake & Allison, 1974) is the idea that access to a contingent response will reinforce an operant response if the schedule causes a reduction in the time spent on the contingent response, unless the operant response increases above its free level.

Restriction and Contingent Effects

Do Premack's results really represent “reinforcement” in the familiar sense? Not necessarily. Look first at the simplest possible theoretical case, where the animal can do only two things, and always does one or the other. That is, there are just two mutually exclusive and exhaustive activities. This situation is illustrated in Figure 2.1, which shows the amounts of time devoted to each activity along the x- and y-axes. If the observation period is fixed, the times of
these two activities must sum to a constant:

\[ x + y = K, \]  

(2.1)

where \( K \) is the duration of the period of observation. **Equation 2.1** is the diagonal line in Figure 2.1. It represents the time-allocation constraint to which these two activities are subject.

Suppose that under free conditions, the animal devotes 30% of its time to activity X and 70% to activity Y. This distribution, termed the free-behavior point, is represented by the filled dot, \( B_0 \), on the time-allocation line. Clearly, by the Premack principle, activity Y is more reinforcing than activity X, so if we require one time unit of X (the instrumental or operant response) for the animal to gain access to one unit of Y (the contingent response or reinforcer), the level of X should increase, an apparent reinforcing effect. Under our assumptions, X will indeed increase (arrow) when a 1:1 contingency is imposed, but is this really a “reinforcing” effect?

The 45° line through the origin shows the constraint imposed by the 1:1 reinforcement schedule (termed the schedule feedback function). Obviously, where only two behaviors are possible, all the degrees of freedom in this situation are exhausted by the two constraints: time-allocation (the diagonal line) and the fixed-ratio (FR) 1 feedback function. Only one point, \( B \), satisfies both constraints. The increase in behavior X represented by the shift of the behavior distribution from \( B_0 \) to \( B \) is forced by the restriction of activity Y caused by linking it 1:1 to X. Thus, the increase in X is an example of a pure restriction effect: not an adaptive response, but merely a forced change.

To obtain something more than a restriction effect under these conditions, more than two activities must be possible. But clearly, even where an adaptive response to the constraint – a contingent effect – is possible, we cannot take a given change at face value. Quite likely some part of the change may be adaptive, but some will also be owing to restriction. There are two ways to distinguish which is which: experimental and theoretical.

**Experimental.** The test for a restriction effect is to ask whether the increase in the operant response associated with the imposition of a contingency is greater than, or merely equal to, the increase produced by just restricting the proportion of time the animal can devote to the contingent activity. The effect of restriction is illustrated in Figure 2.1 by the vertical line through point \( B \). Under these conditions (two mutually exclusive and exhaustive activities) restriction of the contingent response to the level attained under the contingency condition yields the same increase in the operant response, despite the lack of any contingent relation between the two activities. Premack’s early experiments lacked the necessary controls to rule out restriction effects. Nevertheless, later work has borne out the idea that a more probable activity will generally reinforce a less probable one on which it is contingent and that this in-

---

64 The distinction between restriction and contingent effects in behavioral psychology is essentially the same as the distinction in microeconomics between income and substitution effects. For example, if the price of tea rises and people buy less tea, part of the decrease reflects the fact that they have only so much money to allocate to hot drinks and since tea is more expensive they must buy less of it. But part also reflects the fact that coffee is now relatively more attractive and so they spend part of their allocation on coffee rather than tea. The distinction was invented independently in psychology and it took a few years before people recognized that the two are the same.
crease is greater than that expected merely from restriction.

**Theoretical.** Premack drew attention to the fact that there is something special about the way animals allocate their activities under free conditions. Anything that tends to shift the allocation away from the one preferred, such as a 1:1 schedule constraint between a less and a more-frequent activity, leads to a shift in the level of both that tends to restore something close to the previous equilibrium. A number of theories have been proposed to describe how the system adjusts to the constraint imposed by a reinforcement schedule. Typically, they take into account both the equilibrium point, under free conditions, and the displacement forced by the addition of a schedule constraint.

For example, Figure 2.2 shows the possible effects of imposing a FR 1 schedule of response X (operant response) to obtain access to response Y (contingent response). I assume that more than two activities are available (although only two are shown), so that the time-allocation constraint can be neglected. As before, $B_0$ is the free-behavior point (i.e., the preferred distribution of the two activities). The ratio constraint forces a fixed ratio between X, the less-preferred, and Y, the more-preferred activity. $B_0$ does not lie on the constraint line, so the organism will deviate from the preferred level of one or both activities. The two dashed lines show the two extreme adaptations. The vertical line, with equilibrium at point $B_1$, is what happens if activity X continues to occur for its preferred time, so that the constraint forces a large reduction in activity Y. The horizontal line shows what happens if activity X increases enough above its preferred level to prevent any change in activity Y. In the usual operant experiment, activity X is a low-cost activity like lever pressing (rat) or key pecking (pigeon). Y is a high-cost activity such as eating, where reductions are costly for a hungry animal. Hence, most of the change forced by the schedule takes place in activity X; the effect of the schedule is usually a large increase in X, to a point such as $B_3$ in the figure.

For a formal optimality analysis, we need to know two things about each activity to predict the effect of an operant schedule: its free level, $x_0$, and the cost of deviation, $C_x$, from that level:

$$C_x = f(x-x_0),$$

where function $f$ describes how cost increases with deviation. Most models assume that $f$ will be some kind of nonlinear, positively accelerated function (but I’ll show that a much simpler model covers much the same ground). If each activity is independent, the total cost imposed by a given schedule is just the sum of the individual costs, $\Sigma C_i$. Knowing the schedule feedback function; knowing the coordinates of $B_0$; and knowing function $f$, it is usually straightforward to compute the allocation of the operant behavior, $x$, that will minimize total cost. If this allocation is the same whether the schedule involves a contingency between X and Y (e.g., FR 1) or not (i.e., the matched time-constraint schedule, the dashed line in Figure 2.1) then the change predicted in X is purely a restriction effect. But if the two predictions for $x$ differ, the difference is the pure contingent effect.

**Molar versus Molecular.** Implicit in Premack’s analysis is an important assumption: that reinforcing effects depend only on molar measures such as the total proportion of time devoted to an activity or its rate, averaged over a substantial period of time (minutes, hours – the precise averaging period is unspecified, an obvious weakness). He was not at all concerned with molecular properties, such as the pattern in time of particular activities, or contiguity relations between the operant and contingent responses. Many experiments have shown that contiguity – pairing in time – is very important for demonstrating reinforcing effects. Experiments have also shown that reinforcement schedules affect the temporal pattern of an activity, not just its average frequency. So purely molar analysis is incomplete. Nevertheless, molar analysis has been the predominant theoretical approach to free-operant behavior.
Optimal Policy on Ratio and Interval Schedules

If a ratio contingency is imposed that forces behavior away from the free-behavior point, $B_0$, in Figure 2.2, animals typically settle for the sort of compromise represented by point $B_3$: the contingent response, $Y$, decreases somewhat (but not as much as if $X$ had remained at its paired-baseline level, i.e., not to $B_1$), and the operant response, $X$, increases (but not as much as necessary to maintain the level of $Y$ at its paired-baseline level, i.e., not to $B_2$).

When I first thought to diagram reinforcement schedules in this way (Staddon, 1976), the diagram itself immediately suggested a simple hypothesis to explain their effects. Suppose that animals vary their allocation of time among the available behaviors so as to get as close as possible to the free-behavior point, which all theories acknowledge as the animal’s preferred state (economists call it the bliss point – economics may be the dismal science, but it has some cheerful terminology). This prediction is illustrated in Figure 2.3: it just amounts to dropping a perpendicular from $B_0$ to the straight-line ratio-schedule feedback function. The intersection point, $B_3$, is then the predicted value for $x$. The minimum-distance (MD) prediction (Staddon, 1979) always falls somewhere in between the two extreme possibilities. This same calculation can be repeated for a range of other ratio values, X:Y, 2:1, 4:1, etc. The resulting set of $B_3$ values defines a curve, termed the response function. For the simple MD model it is in fact a segment of a circle, whose diameter is the line $O B_0$.

How does this prediction, a circular ratio-schedule response function, compare with data? The answer is, not very well. In most cases, for most ranges of ratio values, for food reinforcement, for rats lever-pressing, for pigeons pecking keys, for open and closed economies the response function is indeed an inverted-U shape, but it is nowhere near circular and the right-hand segment is usually close to a straight line (see reviews in Staddon, 1979, 1983).

There are at least two reasons why the data don’t follow the circle prediction. One is units. In conventional operant conditioning experiments key pecking is measured as pecks per unit time and reinforcement rate as food deliveries per unit time, but there is nothing that makes these two equivalent in terms of time. A peck does not take up as much time as a food reinforcement (which is typically three or four seconds access to a grain hopper). At the very least, the predicted number of pecks needs to be increased to reflect their shorter duration. But even with this correction in the vertical scale, the predicted increase in pecking falls far short of what is observed. On an FR 2 schedule, for example, a hungry pigeon may peck almost twice as fast as on an FR 1, yet the MD model shows it responding at about the same rate, and no alterations of scale materially affect this prediction. What is missing?

The Premack principle, and the simple minimum-distance concept of a bliss point, equates the reinforcing value of an activity with the proportion of time it takes up, a single value. But it’s pretty obvious that the value of an activity to an animal has at least two components:

\[^{65}\] In an “open economy” the animal is chronically food deprived, e.g., at 80% of free-feeding weight, and is exposed to the experiment only for a few hours each day. In closed economies (e.g., Collier, Hirsch & Kanarek, 1977), the animal is not food deprived, is exposed to the schedule every day all day, and must get all its food this way. Closed economics are discussed in more detail in Chapters 8 and 9.
how frequent it is, and how willing the animal is to give it up. There is absolutely no reason these two properties should be correlated. Very infrequent activities, like sex, for example, may be very reinforcing and animals may be very reluctant to forego them, even though they are relatively rare. Conversely, very high-frequency activities, like wheel-running in the rat, may be readily foregone if some attractive alternative is available. In short, the second problem with the simple MD model was that it measured cost only by Euclidean distance (the length of line $B_0B_3$), taking no account of different costs for different activities, which is to say different costs in the $x$ and $y$ directions. With the simple addition of a cost parameter for each activity, the MD model makes predictions much closer to empirical ratio-schedule data.

The Objective Function. Given an independent and a dependent variable (e.g., rate of food delivery, rate of key pecking), the essential ingredients of an operant optimality analysis are the constraints (reinforcement schedule feedback function, time allocation) and a cost function. The first step is to combine constraints and costs in a single expression, termed the objective function. The second step is to find the value for the dependent variable that minimizes the objective function, i.e., yields the smallest total cost (or the greatest total utility, if the objective function is written in terms of utility rather than cost). I illustrate the process first with a very simple response-cost model and then with the MD model.

Ratio Schedules

On a ratio schedule, the average time interval between food deliveries, depends on response rate:

$$I(x) = \frac{M}{x}, \quad (2.3)$$

where $I$ is the average interfood interval, $M$ is the number of responses in the ratio and $x$ is response rate. If the animal responds once every second and $M = 5$, it will get food at 5-s intervals. Equation 2.3 is the objective function for the simplest possible operant optimality model, where the sole cost is the time between food deliveries: time is the currency that is to be optimized — in this case, minimized. The task for the theorist is to discover the value of response rate, $x$, that minimizes average interfood interval. It doesn't take rocket-science to see that $x$ should be as large as possible: the animal should respond as fast as it can. This is also the commonsense conjecture: if getting food as soon as possible is all that matters, then on a ratio schedule you need to respond as fast as you can.

This conclusion is not very helpful because it gives the same answer for any ratio value — no matter what the value of $M$, the optimal policy is to go flat out — and because it puts no limit on response rate. But we know that animals go at different rates on different ratio schedules. What's missing from the analysis? One answer might be that I have ignored the time constraint, but since that affects every activity it doesn't actually alter anything. What about the cost function? An obvious possibility is that Equation 2.3 neglects the energetic cost of key pecking: food delay may be costly, but surely key pecking, especially at the high rate characteristic of ratio schedules (the prediction about this is correct, at least) is also costly. So let's add the assumption that pecking incurs a cost proportional to its rate. This just means adding a term $Qx$ to the objective function, where $Q$ represents the cost of each peck per unit time and $x$ is peck rate. The cost function now has two components: time (interfood interval) and responses. Thus, the new delay-plus-response-cost
Adaptive Dynamics

The objective function, \( C(x) \), is:

\[
C(x) = \frac{M}{x} + Qx. \tag{2.4}
\]

I’ll call this the response-cost (RC) model. It is easy to find the value of \( x \) for which \( C(x) \) is a minimum. Figure 2.4 illustrates it graphically. The Figure shows Equation 2.4 plotted for three values of the ratio value, \( M \): 2, 4 and 8. Notice two things about these curves: (a) The minimum cost occurs at a finite response rate. When each response costs something, it doesn’t pay to go flat out all the time. (b) The response rate that minimizes cost increases with ratio value. (The reason is that \( C(x) \) is made up of two components: \( M/x \), which declines as \( x \) increases, and \( Qx \), which increases with \( x \). The point where the influence of the \( Qx \) component becomes dominant shifts to the right as \( M \) increases.)

Both these features are encouraging. Response rate on ratio schedules is high, but usually less than “flat out.” And empirical results from rats, pigeons and several other species all show that over most of the typical range of ratio values, steady-state response rate does indeed increase with ratio value. Some typical experimental results with ratio schedules are shown in Figure 2.5. The data are from a group of Guinea pigs each pressing a lever for water reinforcement (Hirsch & Collier, 1974, Figure 1). The animals had access to the lever all day and obtained all their water ration via the schedule (closed economy). The figure shows number of responses per day plotted against ratio value. Separate symbols are the data: response rate increases with ratio value, and the points for the 10-s condition are generally higher than for the 20-s condition. The two lines are the predictions of the RC model for the 10- and 2-s conditions and they are in rough accord with the data.

The predictions were derived by finding the minimum of Equation 2.4, with the effective ratio defined as the ratio of water obtained to responses made under each condition (i.e., a ratio of 5 with 20-s access gives an effective ratio of 4-s access per response). Differentiating Equation 2.4 yields:

\[
\frac{dC}{dx} = -Mx^{-2} + Q \tag{2.5}
\]

which, when set equal to zero, yields:

---

66 Since I’m expressing response cost as proportional to rate, why not use rate as a currency to be maximized, rather than time as a currency to be minimized? The model would then become: \( U(x) = x/M - Qx \), where \( U(x) \) is the utility of response rate \( x \). The problem is that this version doesn’t work: it predicts either flat-out responding, or none at all, depending on whether \( Q < 1/M \) or the reverse. This is one of several reasons why time seems to be a better currency than rate. Unfortunately, research on operant behavior has been dominated by the rate variable. (On the other hand, the prediction of this model that animals should cease responding when the ratio value is too large is actually borne out.)
which is the function plotted in Figure 2.5. You might reasonably object that the energetic cost of responding is probably not simply proportional to response rate, very high rates should, perhaps, be disproportionately more costly. Good point, and easy to incorporate into the model. A more general cost function would then be \( C(x) = M/x + Qx^w \), where \( w > 1 \), which yields the optimal solution:

\[
x = \left( \frac{M}{wQ} \right)^{w+1}
\]

which is not very different from Equation 2.6.

**Variable-Interval Schedules**

Exactly the same analysis applies to variable-interval (VI) schedules. The only difference is in the schedule feedback function. How does interfood interval depend on response rate on VI schedules? Figure 2.6 shows that on a VI schedule the average interfood interval is the sum of two delays: the average delay set by the VI schedule timer, \( I \), plus the average delay between setup and the reinforced response, which is determined by the animal's average response rate. If response rate is random in time, this average delay is equal to the reciprocal of the average response rate, i.e., \( 1/x \) (this may seem puzzling; see Box 2.2 for more details). Putting these two together we arrive at the following expression for the (obtained) average interfood interval:

\[
I_{\text{avg}} = I + 1/x
\]

To find the optimal policy, we plug Equation 2.7 into Equation 2.4 in place of \( M/x \), the expression for interfood interval on ratio schedules. The result is

\[
C(x) = I + 1/x + Qx.
\]

Differentiating and setting to zero shows that to minimize this cost function, the animal should respond at a constant rate: \( x = \sqrt{1/Q} \) — as opposed to \( \sqrt{M/Q} \) for the ratio schedule. How do the molar data from VI schedules compare with these predictions?

The relevant data for pigeons, key-pecking for food reinforcement on VI schedules, are well known: the rate of key pecking\(^{67} \) is a negatively accelerated function of obtained reinforcement rate (Catania & Reynolds, 1968; Herrnstein, 1970). This empirical function is a very reliable finding for both rats and pigeons on VI schedules. But the function is in fact approximately constant over quite a wide range, as the RC model implies. Another prediction: because response rate on VI is predicted to be \( \sqrt{1/Q} \) vs. \( \sqrt{M/Q} \) on a ratio schedule, it’s obvious that for any ratio value greater than one, response rate should always be higher on the ratio. This is one of the oldest and most reliable findings about ratio and interval schedules. Several careful comparisons have been made (e.g., Catania, Matthews, Silverman & Yohalem, 1977; Lea & Tarpy, 1982). Conclusion: The RC prediction for molar response rates on ratio and VI schedules is pretty good.

The minimum-distance model (Figure 2.3) looks very different from the RC model, but makes similar predictions. Recall the basic assumption: that the animal minimizes the deviation \( B_0B_3 \) in Figure 2.3. \( B_0B_3 \) is just the square root of the sum of the squares of the two other sides of

---

\(^{67}\) The time taken up eating from the feeder is excluded from these calculations.
the triangle, \( B_0B_1 \) and \( B_1B_2 \), i.e., \((x_1-x_0)^2 + (y_1-y_0)^2\), where \(x_1\) and \(y_1\) are the coordinates of point \(B_1\) and \(x_0\) and \(y_0\) are the coordinates of \(B_0\). If we minimize the square root of something we also minimize the thing itself. Thus, the cost function for the simple minimum-distance model is just

\[
C(x) = (x_0-x)^2 + (y_0-y)^2.
\]

If we now add a parameter, \(c\), to reflect the greater importance of regulating activity \(Y\) (say, if \(Y\) is eating and \(X\) key pecking) we arrive at the formula

\[
C(x) = (x_0-x)^2 + c(y_0-y)^2, \quad c > 1,
\]

which is the cost function for the MD model. Parameter \(c\) is termed the cost-of-deviation (CoD) for activity \(Y\). To get the objective function for ratio responding, we need to substitute the ratio feedback function for \(y\):

\[
C(x) = (x_0-x)^2 + c(y_0/x)^2,
\]

if we assume that the preferred level for the operant response is zero, this expression reduces to

\[
x^2 + cy_0^2 - 2cy_0M/x + cM^2/x^2.
\]

Differentiating and setting to zero yields

\[
\frac{dC}{dx} = 2x + \frac{2cy_0M}{x^2} - \frac{cM^2}{x^3} = 0.
\]

To derive the full objective function for the MD model, we need to incorporate the constraints imposed by time-allocation as well as the feedback function and. Changing our notation to allow for an indefinite number of activities, let \(x_i\) be the fraction of time taken up by \(i\)th activity, then the time-allocation constraint is just

\[
\sum x_i = 1, \quad (2.12)
\]

where \(x_i\) is the fraction of time taken up by each activity. The feedback function, \(y = F(x)\), will of course be different for different schedules. Using the same notation, the cost function becomes,

\[
C(x_i) = \sum c_i (x_0 - x_i)^2, \quad (2.13)
\]

where \(c_i\) is the CoD for the \(i\)th activity. To obtain the minimum cost value for the operant response, we set up the Lagrangean, \(F(x_1, \ldots, x_N; \lambda_1, \lambda_2)\) take partial derivatives and solve in standard fashion (cf. Glaister, 1972; Staddon, 1979, Appendix A). The result is a messy implicit expression that shows the rate of the operant response as a function of the rate of the contingent response.

There are three features about the MD objective function that are different from the simple RC model. First, it is a two-parameter model. Each activity is characterized by two parameters, its paired-baseline level (e.g., \(y_0\)) and its cost-of-deviation (e.g., \(c_i\)). The RC model has only one parameter, \(Q\), which represents response cost per unit time. Second, the MD model is based on the rates of activities, rather than their delays. And third, it is nonlinear in the sense that it assumes that the cost of a given deviation increases as the square of the deviation. The latter makes intuitive sense. A drop in feeding rate of 10 gm/hour will obviously be more costly to the inhabitant of a prison camp, starving on a subsistence diet, than to a well-fed suburbanite. The further the feeding rate from \(B_0\), the more costly additional deviations become. This idea of positively accelerated cost (which is equivalent to negatively accelerated value) has interesting implications for choice behavior.

The objective function with CoD (cost-of-deviation) parameter solves the problems with the MD model that we identified earlier. If \(c\) is much larger than unity, an MD animal will freely increase the rate at which it makes low-CoD response \(X\) so as to maintain approximately con-
stant the level of high-CoD response Y. If c is high enough, an MD animal will almost double its response rate when we increase the FR value from one to two, for example, thus maintaining reinforcement rate almost constant. With these amendments, the MD model makes tolerable predictions of the empirical relations between response and reinforcement rates on variable-interval, variable-ratio and several other reinforcement schedules. In most cases, the predicted relation is an inverted-U: response rate is low at very low and very high reinforcement rates, highest at intermediate rates. The function for ratio schedules is tipped to the right and is everywhere above the function for variable-interval, which is tipped to the right.

The MD model is not particularly simple algebraically, nor does it make strikingly better predictions than some other optimality and economic models. But it is one of the first attempts to show how well-known molar patterns of behavior on reinforcement schedules might be unified through optimality analysis. Since one of the main aims of science is to show how apparent diversity – of elements, of species, of schedule performances – can be reduced to the action of a single process, this was a useful step. The MD model also brought out an important difference between strong reinforcers, like food for a hungry pigeon, and weak reinforcers, like the opportunity to play with a plunger for a Cebus monkey: feeding is a highly regulated activity, whereas play is not. In Premack’s original view, everything of importance about an activity is contained in its free, paired-baseline level: point $B_0$ in Figure 2.1. The MD model showed that activities differ in more than one way: in their free levels, yes, as Premack pointed out; but also in the degree to which they are regulated, indicated by a parameter that represents the different costs of deviation from the free-behavior point. The higher the CoD parameter, the better an activity will be regulated – the harder the animal will work to maintain it at its free level. Strong reinforcers like food and sex may or may not have high operant levels; they certainly have high costs-of-deviation. Weak reinforcers, like the opportunity to run in a wheel, may or may not have low operant levels; they will certainly have low CoD. The MD model also embodies a feature shared with most molar optimality theories of behavioral regulation: a nonlinear cost function. The cost of rate deviations is proportional to the square of the difference. But notice, cost in the RC model is also nonlinear in terms of rate — though linear in terms of time. The two models are not as different as they appear — and the RC model is much simpler.

**Fixed-Interval Schedules**

The RC model also works pretty well on fixed-interval (FI) schedules, but it must be applied in a different way. FI schedules are simpler than VI, because an FI schedule can be fully specified by a single parameter, the interval value, $I$. A VI requires specification of the distribution of intervals, not just their mean. Why should we not apply to FI exactly the same analysis we have applied to VI? The answer is that we might get the right answer (average rate on FI varies little with interval duration, as the RC model predicts), but we would be missing the point. The point is that mammals and birds do not respond at a steady rate on FI schedules, they time the interval. This is one of the most well-known discoveries made by Skinner and his students. After some training the cumulative record of FI performance shows a pause followed by an accelerating rate of responding (Figure 2.7). In a well-trained subject, the time to the first response in each interval (pause or *wait time*) is proportional to the interval duration. This is
Adaptive Dynamics

The Strategy Set. In most operant experiments, the set of activities we need to analyze is fixed by the experimenter, who defines what keys need to be pecked or levers to be pressed. These manipulanda define the set of to-be-reinforced activities. But there are some procedures, like foraging in the open field or timing, where the animal is free to define the set of possibilities itself. Specifying this strategy set in advance, in a non-arbitrary way, poses a substantial theoretical challenge.

On FI schedules, an animal capable of timing can conserve energy and time by learning not to respond until the time reinforcement comes due. It is limited only by its accuracy. Notice in Figure 2.7 how the wait time varies somewhat from interval to interval. The way that variability in wait time imposes a cost is illustrated in Figure 2.8. In the figure, the FI value is 50 and the distribution of wait times is approximately Gaussian, as shown. The animal must trade off wasted pecks (pecks that occur before the FI timer has timed out) against unnecessary delays (wait times longer than the FI value). For example, if it sets the peak of its wait distribution well to the left, at 20, say, it will suffer very few delays, but will waste many pecks. Conversely, if it sets the peak at 80, say, it will waste almost no pecks but will get reinforced on average much later than necessary.

Suppose that the animal responds at a steady rate, $x$, once it has begun (this is a reasonable approximation to FI behavior, particularly for well-trained animals and relatively short – $< 60$-s – intervals). Then the response-cost objective function in this situation looks like this:

$$ C(x, f) = Qx \int_0^I f(t) dt + \int_I^\infty tf(t) dt, \tag{2.14} $$

where $f(t)$ is the distribution of wait times as a function of time; $x$ is the “running” rate, once responding has begun, $I$ is the FI value and $Q$ is the cost of each “wasted” response, scaled in delay units. The first term gives the cost of wasted responses, the second the cost of unnecessary delays. (There is also a constant term that corresponds to the fixed cost imposed by the minimum interreinforcement interval, $I$. But I omit it, since it does not affect the optimality analysis.) The minimum of this function will depend both on the relative cost of responses vs. delays and on the way that $f(t)$ varies as a function of its mean. Notice that the costs are both strictly proportional: $Q$ and running rate $x$ are simple multipliers, and the second term is just the mean delay. This proportionality allows us to draw one conclusion, at least, without knowing the details of $f(t)$: if the shape of the distribution is invariant with respect to its mean – that is, if the mean increases, the standard deviation and other moments increase proportionately so that if the distributions are scaled so that their means coincide, they lie on top of one another – then the optimal solution will also be the same.

In other words, if I were to change the units of the $t$ axis in Figure 2.8, you would not expect the relative position of the optimal solution (vertical line) to change. In fact there is considerable evidence that over quite a wide range of interval values, the distributions of several

---

68 Time devoted to an unnecessary activities entails an opportunity cost, i.e., a cost of activities foregone. The pigeon could be doing something useful instead of pecking fruitlessly for food.

69 Best evidence is that the distribution of FI wait times is Gaussian on a logarithmic, not linear, time scale (Staddon & Higa, 1999; Wynne & Staddon, 1988) but I show the linear case for simplicity. See Chapters 13 and 14 for a fuller discussion of interval timing.
measures of interval timing are roughly proportional, that is: scalar (Gibbon, 1977; Zeiler & Powell, 1994). The RC model implies that when distributions are scalar, measures of timing (temporal control) such as wait time should be proportional to the interval value, which is approximately true. Pigeons should, and usually do, wait twice as long before responding on a FI 120-s schedule as on a 60-s schedule. This is called linear waiting. I will have more to say about both scalar and proportional timing in later chapters.

In sum, the simplest kind of response-cost analysis seems to do a tolerable job of summarizing the main regularities in time- and response-based reinforcement schedules. Are we then converging on a general theory? Can we say that a sort of higher rationality, whose characteristic feature is linear cost of responses and time delays, is the key to all operant behavior? Before reaching a conclusion about the optimality approach to operant behavior, I need to say something about behavioral economics, which is the topic of the next chapter.
Chapter 3

ADAPTIVE FUNCTION, II: BEHAVIORAL ECONOMICS

Behavioral economics and optimality analysis of behavior are one and the same, but each emphasizes different techniques and their historical origins are different. This chapter traces some of the resemblances. It ends with an evaluation of the relative merits of optimality vs. causal models for operant behavior. This survey is highly compressed. If you want more details, check out an introductory economics textbook such as Mankiw and Mankiw (1997) or Samuelson and Nordhaus (1997).

A question at the core of economics is, What is the source of value? Why, for example, does a 100-carat diamond, too large for a jewel, too small for a paperweight, fetch millions of dollars, whereas bread, a staple food, is cheap? Oscar Wilde once remarked that a cynic is a man who knows the price of everything and the value of nothing. Wilde’s witticism suggests one answer to the value question. It was most famously proposed by Scottish proto-economist Adam Smith in his great work An Inquiry into the Nature and Causes of the Wealth of Nations (1776) — but also by his predecessor Richard Cantillon and, even earlier by Catholic philosophers. These folk noticed that what is a defect in a man may be a virtue in an economist. What matters is not value, an abstract unknowable, but price. And price can be explained as the outcome of supply and demand. Smith reasoned as follows:

The market price of every particular commodity is regulated by the proportion between the quantity which is actually brought to market, and the demand of those who are willing to pay the...price....When the quantity of any commodity...falls short of the effectual demand....the market price will rise... (1776/1976, p. 73)

In modern terms, the price of a commodity is explained as an equilibrium between two curves, one of supply the other of demand. The supply curve says how much will be produced at a given price. It is a rising curve — the higher the price, the more will be produced. The demand curve, on the other hand, is falling — the higher the price, the less will be bought. The point where the two curves cross is the market clearing price, the price at which all who wish to sell are able to sell, and all who wish to buy are able to buy.

Of course, we cannot avoid the problem of value if we want to understand why demand and supply curves have the shapes they do. A theory favored by both Adam Smith and the father of communism Karl Marx, points to the inputs necessary to produce a commodity, labor and time. Bread is cheap, the labor-time theory has it, because it is easy to produce; diamonds are dear because they are hard to find. But then so are four-leaf clovers, and they have no market price at all. The labor-time theory is not correct. I discuss some alternatives in a moment.

Demand Curves. Some years ago, several researchers pointed out that behavioral data from animals working on ratio reinforcement schedules conform to the law of demand. As the “price” (work requirement — ratio value) increases, the amount “bought” (reinforcement rate) falls (cf. Lea, 1978). Figure 3.1 shows an example, replotted from a data set I used in Chapter 2. It shows the number of reinforcements per day obtained (“bought”) under different ratio sched-
ules ("prices") in the Hirsch and Collier (1974) experiment. The curve is typical: as the ratio value increases, the amount of reinforcement obtained falls. The "demand law," that as price increases effective demand falls, holds in the world of reinforcement schedules as in the economic world. (There are a few exceptions, however: so-called Giffen goods are bought more, not less, as their price increases. Staple foods, prestige items – and drugs of abuse – are sometimes Giffen goods.)

Demand curves illustrate the regulatory character of responding on ratio schedules. A horizontal demand curve indicates a perfectly regulated activity: no matter what the cost, the amount bought holds constant. If we had to pay for the oxygen we breathe and the price doubled, we would willingly spend twice as much. Demand for oxygen is perfectly inelastic. Demand for anything that is essential and has no substitutes (health care, for example) tends to be inelastic. The data in Figure 3.1 suggest that 100 reinforcements a day – the flat part of the curve – probably corresponds to the minimum water intake necessary for the animals to survive. A downward-sloping demand curve indicates imperfect regulation, elastic demand. Demand for most things — candy, cigarettes, chicken dinners, books — is elastic: when the price rises, the number bought falls. On ratio schedules, complete absence of regulation corresponds to a hyperbolic demand curve: \( R(x) = x/M \), where \( M \) is the ratio value and \( x \) is a constant response rate.

**Labor-Supply Curves.** Reinforcement-schedule data also provide a counterpart to the economist’s labor-supply curve. A plot of response rate against ratio value (e.g., Figure 2.5) shows how the amount of labor supplied (response-rate) depends on the wage rate (ratio value). If ratio value is regarded as a wage rate rather than a price the curve is comparable to what is termed the labor-supply curve. The typical labor-supply curve is “backward-bending” (by the idiosyncratic convention in economics, the \( x \) - and \( y \) -axes are reversed, and the \( x \) -axis reads from right-to-left rather than left-to-right; in our coordinates, the curve would be an inverted-U). The idea is that when the wage rate is low (much work = little reward: high ratio values), people are not willing to work much – the supply of labor (by an individual or a group) will be low. As wage rate rises (ratio value decreases), the amount of labor supplied rises to a maximum; but when the wage rate is very high, labor is withdrawn, because now people have enough money to want to take time off to spend what they have earned.\(^{71}\)

---

\(^{70}\) After Robert Giffen (1837-1910), British statistician and economist.

\(^{71}\) The backward-bending labor-supply curve gave rise to the once-famous “Laffer curve,” which was supposedly sketched on a paper napkin in 1974 by California economist Arthur Laffer as he vigorously argued to a lunch-companion the case for lower income-tax rates. He justified his proposal by the following argument. The effective wage rate depends on the tax rate, since we can only spend what is left after income tax has been extracted. Tax rates are progressive, so that the higher the wage rate the higher the proportion the earner loses in taxes. Hence, there must be a point at which higher taxes begin to depress the total amount of productive work that is done. If the total amount of work done decreases enough, then the total tax "take" will begin to decrease, even though the tax rate is high. The question is: When is the tax rate too high, so high that the depressive effect on total amount of labor supplied exceeds the added revenue brought in by virtue of the high rate? Laffer argued that that time is now, so that taxes ought to be reduced. This argument was persuasive to the Reagan administration and his tax reductions did in fact bring in more money.
A reinforcement-schedule example of labor-supply is shown in Figure 3.2. It shows data from a single pigeon pecking a response key for food reinforcement on different random-ratio schedules, ranging from 12.5 to 400. The data are fairly typical of results from open-economy experiments. The functions are inverted-U shaped: the animals respond at a low rate when the ratio value is very high or very low, and at a high rate at intermediate ratio values.

The data in Figure 3.2 are different from the closed-economy data in Figure 2.5, which show a monotonically increasing response-rate vs. ratio-value function. Only the rising part of the function in Figure 3.2 is regulatory. Why are these curves different? As always when attempting to understand behavioral data, the first thing to look at is the procedure: how, exactly, were these data obtained? The data in Figure 2.5 are from a group of Guinea pigs, in a closed economy (24-hour access to the reinforcer) and water reinforcement. The data in Figure 3.2 are from a pigeon working for food in an open economy (short, daily sessions, supplemented by extra food after the experimental session if necessary). In deciding which of the many procedural differences is responsible for the different results, we have a lot to choose from: species, type of reinforcer, length of session, type of economy, level of food deprivation, individual vs. group.

But the most likely reason for the different results is the different levels of motivation typical of open and closed economies. In a closed economy, the animal must be able to get all the food it needs via the schedule, so animals under such conditions are not excessively hungry. But in open-economy experiments, rats and pigeons are often maintained at 80% of their free-feeding weights, because they are given supplemental food after each experimental session. Consequently, they spend close to 100% of their time working for food during the experiment.

As we’ll see, up-and-down data like those in Figure 3.2 are qualitatively consistent with the MD model, but not with the response-cost model, which implies a monotonic function.

Preference Structure and Indifference-Curve Analysis

Value — utility — is at the heart of economic theory. Psychologists and philosophers differ on whether it makes sense to give numbers to utilities. Some psychologists say you can, most philosophers say you can’t. Economists tend to agree with the philosophers and accordingly have devised a method that allows them to make predictions based only on value relations of “equal” and “greater than”. The method works like this. Consider two goods, such as bread and milk. Even if we can’t give numbers to the utility of a given quantity of milk or bread, everyone agrees that we can almost always equate the value of bundles of goods. For example, 2 quarts of milk and 3 loaves of bread may be judged equal to another bundle with 4 quarts of milk and 2 loaves, in the sense that we are indifferent as to which commodity bundle we get. There will be a whole set of bundles of this sort, differing in the proportions of bread and milk, but the same in that we are indifferent among them. This set defines an individual indifference curve. Defining utilities this way is termed the method of revealed preference (Samuelson, 1965).

An indifference curve represents a set of commodity bundles that are equal in value. Figure 3.3 shows several indifference curves. All points on a curve have the same value, but the indifference curve does not tell us what that value is. Curves can be ordered in terms of value, however, A being preferred to B, B to C, and so on – even if we cannot be sure by how much A

---

72 Not, in fact, a cruel level of food deprivation, since wild-caught pigeons allowed ad lib. food in the laboratory will increase their weight 20-30% (Poling, Nickel, & Ailing, 1990).
is preferred. When bundles of goods are ranked in order of value, they occupy different indifference curves. For example, if we compare 5 quarts of milk and 4 loaves of bread with 4 quarts and 2 loaves we would prefer the first. The 5,4 bundle must occupy a point on a higher indifference curve than the 4,2 bundle. Thus, a set of indifference curves is defined by the relation of value-equality – among bundles on the same curve; and greater (or less) than – for bundles on different curves.

The point of maximum value is represented by the point in the upper right corner of Figure 3.3. The arrow indicates the direction of increasing preference which, for most real commodities is generally in the direction of more of both goods. This known as the axiom of non-satiation: more is always better. The preference structure for bread and milk, or any other pair of goods, can therefore be represented by a sort of contour map of indifference curves, such as those illustrated in Figure 3.3. The analogy to contours is close: like contours, indifference curves cannot cross one another. Like contours, they represent equal values on a third dimension: height, or value. But a set of indifference curves is a contour map lacking an absolute vertical scale: we know where the mountains are and which contour line is above which, but we don’t know how much higher one is than another.

**Constraints.** Knowledge of the preference structure is not by itself sufficient to predict behavior. We also need to know about constraints. Once a constraint is specified, however, the optimal solution is clearly to settle on the highest indifference curve consistent with it; i.e., an equilibrium at the point where the feedback function is tangent to the highest indifference curve. This is illustrated in **Figure 3.4** for the minimum-distance (MD) model for choice between operant responding on the vertical axis and reinforcement on the horizontal axis. The indifference curves are circles centered on point \( B_0 \). Notice that \( B_0 \) is such that the rate of reinforcement is high, and of operant responding, low, which is typical of operant-conditioning procedures with highly deprived animals. Notice also that the indifference curves are closed, a violation of the assumption of non-satiation: more is not always better for behavior allocation. Constraint lines are simply the feedback functions for different ratio schedules. The dashed line is the locus (path) of points traced out by a series of different ratio schedules in this space (the ratio-schedule response function).

Circular or elliptical indifference curves correspond to the MD model described in Chapter 2. They graph the assumption that points equidistant from the free-behavior (bliss) point are all of equal cost. Since equal cost obviously corresponds to equal utility, it makes no difference whether we do our analysis in terms of cost or utility; the indifference curves, and therefore the predictions, will be the same.

The response function traced out in Figure 3.4 is an inverted-U shape. So also is the comparable function relating response rate to ratio value, which is not shown — but you can see that response rate declines at high ratio values (steep feedback lines), rather than continuing to increase. In other words, the minimum-distance model predicts the usual “backward-bending” labor-supply-type relation between response-rate (labor supply) and ratio value (wage rate).

How does the response-cost model compare? We can easily derive indifference curves from the cost function (Equation 2.4),
\[ C(x) = I + Qx, \]
where \( I \) is the interfood interval, \( Q \) is a cost parameter, and \( x \) is the rate of the operant response. Interfood interval is just \( 1/\text{reinforcement rate} \), i.e., \( 1/R(x) \), so that the equation becomes
\[ C(x) = 1/R(x) + Qx. \] (3.1)
To plot an indifference curve, we just pick a value for the cost, say \( C \), set the right hand side of Equation 3.1 equal to it, and rearrange so as to get \( x \) as a function of \( R(x) \). The result is
\[ x = (C - 1/R(x))/Q, \] (3.2)
which we can then go ahead and plot. Each time we change the value of \( C \), we get another indifference curve. Indifference curves for three values of \( C \) are plotted in Figure 3.5. The response function goes through the points at which the three ratio-schedule feedback functions are tangent to the three indifference curves. As I showed earlier, the RC response function slopes downwards to the right, like a demand curve.

Indifference-curve analysis makes it easy to separate the cost or utility function, which is the essence of any optimality model, from the constraints. The set of indifference curves represents the cost function; the constraint lines then define on which indifference curve, and where on the curve, behavior will lie. I have only described ratio-schedule constraints so far, but it is almost as easy to derive predictions for other schedules, such as variable-

interval. Figure 3.6 shows two indifference curves for the RC model that are tangent to two VI feedback functions. Both points of tangency are at the same response rate, in agreement with the optimal VI policy I derived for the RC model in Chapter 2.

Notice that the VI feedback functions in Figure 3.6 have a shape that makes sense: when response rate is very low, reinforcement rate is almost proportional to response rate. But as response rate approaches the maximum reinforcement rate specified by the schedule, reinforcement rate rises more and more slowly, approaching as an asymptote the maximum rate permitted by the VI.

An inverted-U labor-supply function implies indifference curves resembling those of the MD model; a downward-sloping demand-curve function implies indifference curves like those of the RC model. Figure 3.7 shows that the indifference curves for the MD model differ slightly from those of the RC model. The effect is that the MD model predicts a VI-schedule response function that rises over most of the range (response rate at point \( B_2 > B_1 \)), rather than being constant, like the RC model prediction (Figure 3.6). On balance, the MD analysis — and related graphical models of Rachlin and his associates (e.g., Rachlin et al., 1981) — fit experimental results a bit better than the response-cost model. But the MD model is more complicated and the Rachlin et al. models lack an axiomatic basis. The simple assumptions of the RC model, and its ability summarize almost as much data as the others, give it a useful place. I return in a moment to the problem of model uniqueness.
Although we cannot go from indifference curve to the underlying utility functions, it is straightforward to go in the other direction. Given the form of the cost function, we can easily derive the form of indifference curves. So why do we need yet another method for doing the same thing? What is gained by an indifference-curve analysis — other than epistemological purity?

First, the indifference curve method is potentially experimental: it tells us (at least in principle) how to actually measure the organism's preference structure: by equating commodity bundles. The method of revealed preference is the closest that neoclassical economists come to being pure behaviorists. Like reinforcement theorists, they are almost unanimous in agreeing that the only way to assess value is through the individual's revealed preferences, represented by indifference curves. But in fact, neither economists nor psychologists much use the direct method of asking people or animals to equate commodity bundles. Although there is now a small school of experimental economists, the majority do no experiments whatever. And psychologists, for practical reasons, usually do experiments of the type we have already discussed, and then test economic models indirectly.

Second, a set of indifference curves need not follow any particular mathematical form. It can be anything we like, subject to the logical limitation that indifference curves cannot cross (economists add other assumptions, like non-satiation). While some behavior theorists (e.g., Battalio, Kagel, Lea, Staddon, most behavioral ecologists) have favored preference structures derived formally, from assumptions such as minimum distance or response cost, others (e.g., Hursh, Rachlin, many economists) have favored the flexibility of a graphical approach in which indifference curves are simply drawn in a plausible fashion.

Third, because indifference curves represent the cost function directly, they show the real similarities between models. Figures 3.6 and 3.7 show the indifference curves derived from minimum-distance and response-cost. They are obviously very similar, even though the equations that describe the two models look very different.

**Marginal Value and Substitutability**

It is no accident that indifference curves are assumed to be convex to the origin, that is, curving outwards from the region of maximum value. Convexity is a consequence of a very important property of value: diminishing marginal utility (or increasing marginal cost). What it means is that the more you have of something, the less each additional increment is worth (only Giffen goods are exceptions). Conversely, the more you lose of something, the greater the cost of each additional bit of loss.

Diminishing marginal utility is illustrated in Figure 3.8, which shows a negatively accelerated utility function. The vertical lines indicate increments of the good (e.g., food-rate increments). The two curly brackets show increments in utility associated with successive equal increments in the good. The first utility increment, $\alpha$, is greater than the next, $\beta$, and this pattern continues across the whole curve. Thus, going up and to the right from point A (increasing utili-
ty), the curve shows diminishing marginal utility; going down and to the left from A, the same curve shows increasing marginal cost.

Diminishing marginal utility implies diverse resource allocation – partial (rather than exclusive) preference. For example, given $100 to spend on any mixture of three commodities, coke, beer and French fries, maximum satisfaction is assured if we are indifferent on whether to spend our last dime on more coke, beer or fries. Indifference means that we have equated the marginal values of the three commodities. If each of these goods is subject to diminishing marginal utility, maximization will assure a balanced diet. We will choose some of all three, rather than spending the whole $100 on one.

The relation between partial preference, diminishing marginal utility, and convex indifference curves, is illustrated in Figure 3.9. The figure shows a convex indifference curve for two goods, A and B, each of which shows diminishing marginal utility. The diagonal line represents what economists call the budget constraint: a fixed sum of money, which must be spent on A and B. The slope of the line represents the relative prices of A and B. If A and B are activities, measured as proportions of time spent, the budget constraint corresponds to the time-allocation constraint discussed in Chapter 2. The optimal allocation of resources is at point \( B_1 \), where the constraint line is tangent to the indifference curve. \( B_1 \) obviously represents a partial preference.

What will be the form of indifference curve when utilities do not show diminishing returns? If the utility of each activity is simply proportional to the amount of activity: \( U_1 = aT_1 \) and \( U_2 = bT_2 \), where \( U_1 \) and \( U_2 \) are the utilities of investments, \( T_1 \) and \( T_2 \) are the times invested, and \( a \) and \( b \) are the ratios of utility to time spent. To derive the indifference curve, we pick a total utility (or cost), call it \( C \), and set the combined utilities equal to it, so \( C = U_1 + U_2 = aT_1 + bT_2 \); rearranging gives us \( T_1 = (C-bT_2)/a \), which is a straight line. There is obviously no way that a straight-line budget constraint can be tangent to a straight-line indifference curve. Hence, the optimal policy is to choose activity 1 or activity 2 exclusively. Note that a straight-line indifference curve need not imply that the goods it relates are not subject to diminishing marginal utility. It may mean that the two goods are not independent, so that consumption of one causes satiation for both.

Economists have names for these different kinds of indifference curves: Two goods related by a straight-line indifference curve, where the optimal policy is exclusive choice, are termed perfect substitutes. Examples are two equivalent brands of gasoline: you just pick the one with the lowest price. Two goods related by a convex indifference curve are termed partial substitutes, because a change in price (slope of the budget line) causes a change in preference that falls short of a complete switch from one to the other. Coffee and tea are partial substitutes: if tea becomes more expensive, you may drink more coffee. Two goods related by an extreme convex indifference curve (i.e., a corner, made up of a vertical and a horizontal line) are termed complements. The idea here is that there is a preferred ratio of the two goods (e.g., equal numbers of left and right shoes), so that we are indifferent to additional quantities of either good: a bundle of 12 left shoes and 10 right has the same value as a 10:10 bundle.

These names are not very helpful for analyzing behavior because the terms say more than they mean. Substitution, for example, implies functional equivalence (i.e., that consumption of one good satiates for both), but all it means is a particular form for the indifference curve. As we have seen, a given indifference curve may be generated in several ways, only some of which correspond to functional equivalence.

**Uniqueness of Models**

Any consistent set of behavioral data can be fit by some optimality model. The animal is always (apparently) minimizing something. Even the laws of Newtonian physics, the very model of causation, can be expressed as an optimality model, as the Irish mathematician Sir William...
Rowan Hamilton showed more than 150 years ago (Conway & Synge, 1931). Snell’s law of refraction, for example, can be deduced from a principle of minimum time: given that light travels at different speeds in different media, the angle of refraction (from air to glass, say) is such as will allow a ray to travel from one point to another on the other side of the boundary as quickly as possible. Snell’s law can also be deduced from the principle of wave interference and from quantum electrodynamics (Feynman, 1985). Conclusion: optimality, as a principle, cannot be disproved. But particular optimality hypotheses, such as the idea that animals maximize the rate of reinforcement, or the profitability (yield/time-to-consume) of a food source, can be tested.

Optimization is always subject to constraints. Failure of a particular model can therefore always be attributed either to a deficiency in the model or failure to set the proper constraints. Constraints may be external, such as a reinforcement schedule, or internal, such as limitations on memory or speed of processing or the types of possible operant response. Internal constraints reflect properties of the causal mechanisms involved in operant behavior.

Optimality models are useful, therefore, only when constraints are well understood and when the currency, the thing to be optimized (reinforcement rate, for example), is either fixed for all situations or, if not fixed, varies from situation to situation according to a well-defined law. Existing optimality models for operant behavior do not meet these conditions. Microeconomic models of human behavior, a vast and vigorous field of inquiry, do if anything even worse, because we understand less about the constraints on human behavior than we do about the limits on the behavior of animals, because the motivations of humans are infinitely more varied than the motivations of animals and because experiments on human economic behavior are usually impossible. Lacking experimental test, theory can flourish.

### Implications for Choice

A two-armed bandit is just a two-choice version of the familiar Las Vegas one-armed bandit, that is, a situation in which the animal has two choices (levers, response keys), each of which delivers reward on a probabilistic (i.e., random-ratio) schedule. For example, the subject may be a rat responding on one lever for a slug of Cherry Cola, and on another either for the same thing (Cherry Cola) or for something different (Tom Collins mix). Suppose the two random-ratio schedules are the same, e.g., 25, so that the rat gets access to either reinforcer with probability 0.04 for pressing the appropriate lever. What should it do if both levers produce Cherry Cola? Well, it really doesn't matter, so that the chances are that after much experience, the rat will just develop a position preference and respond exclusively on one or other lever.

These outcomes are just what we would expect with two reinforcers that are perfect substitutes. If the animal has a fixed number of responses to expend, then the constraint line is just like the budget line in Figure 3.9: $x$ responses on the Left means $N-x$ on the Right, where $N$ is the total permitted, and the prices (ratio values) are equal. If the two reinforcers are perfect substitutes, then the indifference curve is also a straight line, so that the (optimality) prediction will always be exclusive choice of one option or the other.

What will happen if the reinforcers for the two responses are different (Cherry Cola and Tom Collins mix)? Now the indifference curve may be convex rather than linear – complete satiation on Cherry Cola may leave the rat still with some appetite for Tom Collins mix, and vice versa (full-up on the entree, you may still have room for dessert). The prediction for equal ratio schedules is now very different: partial preference, rather than exclusive choice. Moreover, if the partial preference favors Cherry Cola, say, then we can increase the ratio on that side without abolishing responding – indeed, the change in preference for a given change in “cost” allows us to estimate the indifference curve directly. The behavior of real rats is consistent with this simple economic analysis (Rachlin et al., 1981).
Now let's look at different food-reinforcer *amounts* for each choice, with equal ratio values for each. No matter what the form of the utility curve for food amount, so long as more is preferable to less, the rational rat should always pick the large-amount option exclusively, and real rats usually do. But we can see the effect of the utility function if we change the procedure slightly: on the Left, the animal continues to get a small amount, say a 40 mg food pellet. But on the Right we give him either a very small (20 mg), or a large (60 mg) pellet, with equal probability. So now 25 responses on the left on average buys our rat 40 mg of food; 25 responses on the right buys him equiprobably 20 or 60 mg, which averages out to the same amount. The question is, which should he prefer? The answer (from the point of view of optimality analysis) depends on how the utility of food depends on its amount. Look again at the negatively accelerated utility curve in Figure 3.8. How much utility (as opposed to how much food) does the rat gain from each of the two options I have described? On the left, the rat gets amount A in the Figure, about 0.8 on the utility axis. On the right he gets either 0.8-α, or 0.8+β. But since α is plainly always greater than β (given a negatively accelerated utility function), then the average of 0.8-α and 0.8+β must be less than 0.8 – so the rat should prefer the side with the fixed food amount to the side with variable amounts with the same average. And real rats as well as other species, usually do (Staddon & Reid, 1987). This aversion to variability is termed *risk aversion*, and it is almost a universal result, whether the choosers are people or animals. The diminishing-marginal-utility explanation for risk aversion is an old one, first offered by the Swiss mathematician Daniel Bernoulli in 1738.

Risk aversion can easily be demonstrated with human subjects. Daniel Kahneman and Amos Tversky (e.g., 1979) have become famous for a series of experiments in which they asked college students deceptively simple questions about decisions involving gambles – and got surprisingly dumb answers. In one experiment, students were asked to decide between the following two outcomes:

*Which of the following would you prefer: $3000 for sure, or a 0.8 chance of $4000?*

Since 0.8 x 4000 = 3200, the gamble has the higher expected value. Nevertheless, the great majority of subjects opted for the sure thing. They were risk averse, as implied by a negatively accelerated utility function.

Risk aversion is in fact rational, in general. It makes good adaptive sense, for at least two reasons. First, the future is uncertain: “a bird in the hand is worth two in the bush.” So, a small “sure thing” now is worth a larger “possible” off in the future. Second, the value of reinforcements such as food depends not just on their amount but also on their distribution in time. An individual may subsist on 100 kg of food per year, say, but he is unlikely to trade a regimen of 0.3kg/day (total: 109.5 kg), for 200 kg delivered at year’s end. A large food reward should be less valuable than two half-size rewards delivered twice as often, although the difference might well be small if the amounts and delays are also small. Since any real-world gamble involves a change in temporal distribution, as well as distribution of amounts, risk aversion will often be adaptive.

**HOW VALID IS OPTIMALITY THEORY?**

Most optimality accounts of operant behavior share two fundamental assumptions: (a) That behavior is sensitive to marginal changes in reinforcement rate; and (b) that animals actively seek to maximize reinforcement rate. How valid are these assumptions?
Is Behavior Directly Sensitive to Marginal Molar Changes?

Ettinger, Reid and Staddon (1987) carried out a direct test of the marginal-sensitivity assumption. They chose a schedule that has linear molar feedback functions, because many optimality models predict a particularly simple adaption to such schedules: a straight-line response function. These models all predict that a change in the slope of the molar feedback function (i.e., a change in its marginal rate of return) should always cause some change in the slope of the response function. The models differ on the direction of the change, but all predict some change.

Ettinger et al. used interlocking schedules, which are combination interval-ratio schedules. If the animal does nothing, the schedule is effectively a fixed-interval. But, rather like a ratio schedule, interfood time is reduced by every response. For example, if the animal does nothing, food becomes available for the first response after 60 s, say. But if the animal makes one response before the 60 s mark, then food becomes available after 55 s; if he makes two, food is available after 50 s, and so on. Thus, the interfood interval is determined according to the formula 

\[ I = T - am \]

where \( m \) is the number of responses that are actually made (excluding the response that actually procures the reinforcer), \( T \) is the minimum interfood interval if no responses are made, and \( a \) is a parameter that says how much each response reduces the time to food.

The molar feedback function for this schedule is obtained as follows. Take a single interfood interval as the unit of analysis. Within that interval, response rate (again, excluding the reinforced response) is just \( x = m/I \); reinforcement rate, \( R(x) = 1/I \). Replacing and rearranging yields the feedback function

\[ R(x) = 1/T + (a/T)x. \] (3.3)

In other words, the feedback function is a straight line whose slope is determined by schedule parameters \( a \) and \( T \).

Figure 3.10 shows the results of one of the experiments by Ettinger et al. The figure shows feedback functions for two sets of interlocking schedules that differed in the slope parameter, \( a \) (two sets of positive-slope lines). Within each group, the slopes were the same but the intercepts \( (1/T) \) differed. Response functions predicted from the MD model are sketched in as dashed lines. Other optimality models predict different changes in the slope of response function. Yet the results are clear: (a) the obtained response functions are indeed approximately linear, but (b) there seems in fact to be only one linear function, which is the same for both sets of schedules. All the data points seem to lie on the same straight line with negative slope. The difference in molar marginal rate of reinforcement between the two series evidently had no effect on response-function slope.

The strong suggestion of Ettinger et al.’s results is that changes in molar feedback functions have no direct effect on responding. Differences between ratio and interval schedules (for example) may in fact be the outcome of local, molecular processes, even though molar accounts often provide surprisingly accurate summaries of the different performances generated by different schedules. The conclusion that molar variables do not have direct effects is reinforced by data from a wide range of other operant procedures, some of which I discuss in later chapters (see Dragoi & Staddon, 1998, for a review).
Do Animals Maximize Reinforcement Rate?

A second assumption of most optimality models is that animals will act so as to maximize reinforcement rate. When they fail to do so, “cognitive constraint” is the usual explanation. For example, on so-called spaced-responding or “differential reinforcement of low rate” (DRL) schedules, subjects are required to space their responses at least \( T \) s apart to get reinforcement. Interresponse times (IRTs) that are too short reset the schedule clock so the animal must wait again. Pigeons and rats have difficulty with DRL schedules if the waiting time is greater than 20-30 s, even though under other conditions, like long FI schedules, they have no difficulty waiting for periods of minutes (Richelle & Lejeune, 1980; Staddon, 1965). Animals treat long DRL schedules more or less as if they were VI schedules, and get reinforced rarely, rather than once every \( T \) s, as the schedule would permit, given appropriate waiting. Various explanations have been offered for this failure, but most of them invoke either some limitation on response inhibition (waiting) or limits on the animal’s ability to remember the time of its last response.

Perhaps a cognitive-constraint account is correct here. But there are simple situations where such an account does not seem to work.

Let me describe a striking example from some experiments with Clive Wynne (Wynne & Staddon, 1988). The experiments used a schedule called response-initiated delay (RID)\(^73\), diagrammed in Figure 3.11. After food reinforcement, the pigeon waits a time \( t \), which is entirely under its control. After the first peck, a second time, \( T \), under the control of the schedule, elapses until food is again delivered (without necessity for further responding). A stimulus change usually signals the change from time \( t \) to time \( T \) (for example, the response key may be red until the first peck and green thereafter), though this does not seem to be essential. The dependent variable is waiting time: \( t \). The independent variables are time \( T \) and possible scheduled relations between \( T \) and \( t \).

Experiment 2 compared three conditions:

- **clamped**: \( T + t = K_1 \), a constant (an FI-like schedule)

- **fixed**: \( T = K_2 \), a constant

- **dependent**: \( T = K_3/(t+1) \), that is, time \( T \) in the current interval is inversely related to the value of waiting time, \( t \), in that interval, and \( K_3 \) is a parameter.

We chose these three conditions because each implies a different optimal strategy, i.e., strategy that gives the shortest interfood interval. There is no optimal strategy for the clamped condition, other than to respond once before time \( K_1 \) has elapsed. We expected that the pigeons would treat the clamped procedure like a fixed interval, with wait time settling at some fixed fraction of \( K_1 \). The optimal strategy for the fixed condition is very different, however: since time \( T \) is fixed, the optimal strategy is for the animal to set waiting time, \( t \), as short as possible, for all \( T \) values. The optimal strategy for the dependent condition is to set \( t = \sqrt{K_3 - 1} \).

Pigeons were exposed successively to all these procedures and several different \( T \) values. The critical test is to look at the relation between \( t \) and \( T \). Are the pigeons sensitive to the programmed relation between \( T \) and \( t \)? Figure 3.12 shows the data. It’s a complex figure, but the main point is clear: irrespective of condition, the relation between \( t \) and \( T \) is the same. In other words, the pigeons seemed to treat all these procedures like a fixed interval schedule of duration \( t + T \). The results for the fixed condition (vertical lines) are particularly striking, because the op-

---

\(^73\) Also termed a conjunctive FR 1 FT (fixed-time) schedule (Shull, 1970; Staddon & Frank, 1975)
Optimal policy is for $t$ to be as short as possible, and independent of $T$. Instead, $t$ values in this condition were proportional to $T$, just as they were in the other conditions. The pigeons were not maximizing reinforcement rate in this experiment and their limitation does not seem to be cognitive, since all that was required in the fixed condition was to respond immediately after food. Instead, their behavior seems to be driven by some kind of obligatory timing rule that sets wait time to be a fixed fraction of the prevailing interfood interval.

The simplest such rule is termed linear waiting, the idea that waiting time tends toward a linear function of the preceding interfood interval. This rule can be expressed formally as follows (Wynne & Staddon, 1988): Let $E$ be the animal’s estimate of the interfood interval. Then, after each food delivery, we assume that the prevailing wait time is changed in the direction of the steady-state linear relation $(aE+b)$ between $E$ and $t$:

$$\Delta t = (aE+b) - t_{N-1}, \quad (3.4)$$

where $\Delta t$ is the change in $t$, $a$ and $b$ are constant and the subscript, $N$, refers to the Nth interfood interval. How is $E$, the animal’s estimate of the interfood interval, derived? The simplest possibility is that

$$E_N = I_{N-1} = t_{N-1} + T_{N-1} \quad (3.5)$$

i.e., the animal’s estimate of IFI duration is just the duration of the previous interfood interval. This idea seems improbable because it implies almost instantaneous temporal learning, while most data seem to show that exposure to tens or even hundreds of to-be-timed intervals is usually required before wait time or some other dependent measure adapts to the interval. But in fact direct experimental tests (e.g., Staddon, Wynne & Higa, 1991; review in Higa & Staddon, 1997) have shown it to be true under many conditions. So, let’s see where the idea leads us. Combining Equations 3.4 and 3.5 yields

$$\Delta t = a(t_{N-1} + T_{N-1}) + b - t_{N-1}, \quad (3.6)$$

which has a stable equilibrium ($\Delta t = 0$) when

$$\hat{t} = a(\hat{t} + \hat{T}) + b. \quad (3.7)$$

Thus, this kind of quasi-dynamic process is sufficient to get the linear equilibrium relation shown in Figure 3.12. How might the idea be further tested?

There is a simple method, that can be explained as follows. First, we can neglect constant $b$ in Equations 3.4-7, because it is usually small, so that **Equation 3.6** becomes $\Delta t = t_{N-1} - t_{N-1} = a(t_{N-1} + T_{N-1}) - t_{N-1}$, which reduces to

$$t_N = at_{N-1} + aT_{N-1} \quad (3.8)$$

Now, suppose we program the schedule so that the delay, $T$, in interval $N$ is proportional to the waiting time, $t$, in the preceding interval:

$$T_N = \alpha t_{N-1} \quad (3.9)$$

where $\alpha$ is a constant. Combining Equations 3.8 and 3.9 yields
$t_N = at_{N+1} + \alpha at_{N-2}.$  

(3.10)

The stability of Equation 3.10 (i.e., whether the process converges to a finite value or diverges to $\infty$) depends on the relative values of $a$ and $\alpha$. The second term is critical. If the weight $\alpha a$ is greater than unity, this term grows without limit as $N$ increases; but if it is less than unity (and given that $a$ is always less than unity – typical values are $a \sim .3$) $t_\infty \to 0$. In short, when $\alpha$ is less than $1/a$, $t$ should converge on zero, and when $\alpha$ is greater than $1/a$, $t$ should increase without limit. These are testable predictions.

**Figure 3.13** shows the results of one pigeon in the appropriate experiment. Capital letters denote large values for $\alpha$ predicted to cause divergence; small letters small values predicted to cause wait time to fall to zero. The results for this animal and three others were more or less as predicted: when $\alpha$ was large, wait times got longer and longer within the experimental session. When $\alpha$ was small, wait times rapidly declined to small values. The experiment is not sufficient to define exactly the rule the animals were following. The results are consistent with a one-back learning rule, like Equation 3.5, but probably also with some rules that allow for an effect of earlier intervals. But the results do show what is wrong with optimality models. Even when behavior appears to be optimal, the animal is not following an explicit optimizing strategy. It is following a rule that may yield optimal behavior under some conditions, as linear waiting does on FI schedules, but will fail under others. Linear-waiting-type rules fail dramatically on schedules such as the fixed condition in the Wynne and Staddon experiment, for example. I conclude that scientific psychology cannot settle for the study of adaptation or optimality alone.

**CONCLUSION: OPTIMALITY, ECONOMICS AND THE LAWS OF BEHAVIOR**

What can we conclude about the optimality approach to operant behavior? First, no optimality model works in every situation. Animals are rarely, if ever, “literal optimizers”; they don't remember the average payoff associated with a given pattern of responding, compare it with the payoffs for other patterns (from a well-defined set of possibilities) and then pick the best pattern – as some views of optimal responding seem often to imply. Under most conditions, people don't behave in this way either. Moreover, the data from the Ettinger et al. and Wynne and Stad-
Adaptive Dynamics

Staddon

don experiments, as well as many others (some are reviewed in Staddon, 1987) strongly suggest that the most important causal relations are not at the molar level at all. Nevertheless, the simple local rules animals often use perform remarkably well in a wide range of situations. Hence, molar optimality analyses also do rather well.

Optimality analyses have revealed unsuspected relationships between implicit choice – the allocation of behavior under relatively free conditions, and explicit choice, between well-defined alternatives. Optimality models still provide the simplest account of risky choice and of choice between partially substitutable reinforcers. Optimality models agree that reward value is a negatively accelerated function of reward amount, so that doubling amount increases value by a factor less than two. In short, optimality models provide a tolerably good picture of what animals manage to achieve, but a poor picture of how they actually do it. Animals often behave optimally; they almost never optimize in any literal sense (Staddon & Hinson, 1983; Staddon, ).

Behavioral economics is just optimality theory by another name, and with a slightly different bag of theoretical tricks: indifference curves instead of objective and cost functions. But the rationale, the results, and the limitations of behavioral economics are exactly the same as those of optimality. Only the names and connotations are different. “Labor supply” and “demand curve” suggest different things than “response cost” and “minimum distance” but the principles behind them are the same.

Specific optimality models always have their limitations. But the idea of adaptation itself – fit to an ecological niche, Darwinian fitness – has a use I have not really touched on in this discussion of the formalism of optimality theory: it gives hints, sometimes very strong and useful hints, about the environmental variables that are likely to be important in causal models. It is no accident that the rate of food delivery plays a key role in many optimality models of operant behavior, for example. The events that act as reinforcers are obviously important to Darwinian fitness. Activities like exploration and play are both reinforcing and adaptive. Not everything that is adaptive is reinforcing, but almost everything that is reinforcing is also adaptive – and apparent exceptions, like addiction to drugs or fatty foods, can reveal either properties of the underlying physiology or properties of the selection environment (i.e., the environment in which these mechanisms evolved). Some drugs seem to affect directly brain circuits involved in food reinforcement, for example; experiments like the one with the response-initiated-delay schedule discussed above, remind us that response-initiated delays are not a feature of any natural environment. And our almost limitless love for calorie-rich foods suggests that our ancestors rarely had the opportunity for tasty feasts and thus little need to show self-restraint.

Late-Victorian biologists, like many contemporary behavior theorists, were much concerned with the discovery of “laws.” Their laws are not like Newton’s law of gravitation, which is invariably true under given conditions. Their laws were not the sort that can be verified by experiment. Principles like Bergmann’s Rule, Cope’s Law and the Law of Conditions of Existence are descriptive generalizations, born of natural history and a search for order. Nevertheless, they provide grist for the theoretical mill. In Chapter 6, “Difficulties of the Theory” in the Origin of Species, Darwin discusses the law of Conditions of Existence, which is an old version of the principle of adaptation. “The expression of conditions of existence, so often insisted on by the illustrious Cuvier, is fully embraced by the principle of natural selection. For natural selection acts by either now adapting the varying parts of each being to its organic and inorganic conditions of life; or by having adapted them during past periods of time...(1872)” Expressed in modern language, Darwin’s point is that adaptation depends on mechanisms. The mechanisms may be phylogenetic, like natural selection and a handful of other principles to do with growth and heredity; or they may be ontogenetic, to do with the mechanisms of learning. The “conditions of existence” exert no effect on their own, there is no literal optimizing adaptation, but only adaptation through the action of causal mechanisms, which may fail when condition diverge substantially from the selection environment. Darwin’s conclusion is also mine: when behavior is
optimal it depends on processes that permit the animal to do things that maximize its fitness. But when it is not optimal, it depends on exactly the same processes. Hence, processes are fundamental, optimization is not.
Orb-web spiders build efficient nets for catching flying insects. How did they develop the design and the program for executing it? No one has seen individual spiders trying out different types of web, as a human designer might do. Spiders don't learn how to weave good webs (although they necessarily adapt each web to a particular environment). Nevertheless, the answer is trial-and-error, but not in the life of the individual, in the life of the species. Ancestral spiders built webs with varying efficiencies. Those that built best were better nourished and had more offspring (i.e., had higher Darwinian fitness). Since individual differences in web-building ability are partly inherited, the ability to build better webs evolved by selection of spiders who made better webs. Web-building, like all adaptive behavior, depends on history. Like all instinctive behavior, it depends more on the history of the species than on the history of the individual.

But living organisms also learn, and the simplest kind of learning is also trial-and-error.

**Trial-and-Error in Plants.**

Green plants feed on sunlight, so it's important for any green plant to get a place in the sun. The chief obstacles are other plants with similar requirements. Where the need for light is not outweighed by other considerations, such as deterrence of foragers or extremes of temperature or wind, plants grow vertically and seek the highest, best-lighted point.

Darwin studied growing plants of several varieties and noticed that the growing tip always moves as it grows (another example of Darwin’s extraordinary powers of observation, in days long before time-lapse photography). He described the movement as follows:

> When the shoot of a hop (*Humulus lupulus*) rises from the ground, the two or three first-formed, whilst very young, may be seen to bend to one side and to travel slowly round towards all points of the compass. . . From seven observations made during August . . . the average rate during hot weather and during the day is 2 hrs. 8 m. for each revolution . . . . The revolving movement continues as long as the plant continues to grow; but each separate internode, as it becomes old, ceases to move. (Darwin, 1875, p. 2-3)

Two things are happening here, rotation (termed circumnutation) and upward growth (negative geotropism). Given a supporting environment, these two processes, plus a third, phototropism, (growing towards light) help the plant to find the light. All are fixed, instinctive, but taken together they comprise a simple search process. Upward growth, unimpeded, will lead the plant to light. But if a stalk or stick is nearby, the turning plant will twine around it, attaining greater height than it could unaided. It is important that the turning be not completely smooth or predictable (see Figure 4.1): if the shoot misses a support on one turn, it may nevertheless grasp it on the next. If it gets caught in a crevice, continued movement (together with phototropism) will usually allow it to escape.

This example shows how all adaptive behavior grows out of the interplay between processes of variation and selection. Circumnutation plays the role of variation: the plant tip sweeps out an arc “searching” for a bright spot of sky or some vertical support. The environment, in the
form of sticks, stalks and patches of light, then plays selector, blocking the moving tip and constraining it to a particular place. In most environments, the three component processes — photo- and geotropism and circumnutation — work harmoniously together, so that the plant tends to improve its situation.

**Kineses**

A note in *Nature* described an extraordinary predator-prey interaction — involving bacteria. Wenyuan Shi and David Zusman describe how *Myxococcus xanthus*, “a Gram-negative rod-shaped fruiting bacterium found in damp soils...[which] form[s] large multicellular communities” preys on another bacterium, the familiar *Escherichia coli*. The problem is, *M. xanthus* is a tortoise and *E. coli* is a very fast hare. They write that on some media “*E. coli* swims ... 50-200 times faster than *M. xanthus* ... we became concerned with how slow-moving *M. xanthus* cells can catch fast-moving *E. coli* cells.” How is it that slow *M. xanthus* is able to prey on fast *E. coli*?

Finding the proper habitat, a place not too hot or cold, or too acid or alkaline, safe from predators and with a supply of food, is a problem that must be solved by even the simplest organisms. Protists — bacteria, unicellular organisms of all sorts — possess no nervous system and only the most rudimentary chemical, temperature and vibration sensors. They have no *distance receptors*, sensory systems such as vision or audition able to register the direction and distance of a predator or source of food. Hence, they must avoid bad places and seek good ones using very simple means. The processes these organisms use to find their way about show the properties of adaptive behavior in their starkest form.

In the 1880's the German biologists Engelmann and Pfeffer discovered bacterial *chemotaxis*: the ability of bacteria such as *Salmonella* (the agent of common food poisoning) to move up or down a chemical gradient. (Chemotaxis also plays benign role for humans, because white blood cells use it to find foci of infection in the body.) Pfeffer did the simplest possible experiment. He dipped a capillary tube containing a suspension of food into water containing bacteria. He found that more bacteria entered the capillary than would be expected if they simply moved randomly: after a while, they clustered around the pipette tip (**Figure 4.2**). This is extraordinary behavior for an organism only about two millionths of a meter long, with only the most primitive sensory and motor apparatus, and no nervous system at all.

How do the bacteria find their way to the food? The only information available is chemical concentration. They have no eyes to see, and even if they had, they could not see the source of food directly. They have to use *local* information about the chemical gradient. There are only two ways to detect a chemical gradient: either simultaneous or successive comparison. Simultaneous comparison is the easiest to understand. Given two receptors spaced some distance apart, an animal can move up a gradient by moving ahead and, at the same time, turning towards the most-stimulated receptor. When both receptors are equally stimulated the animal will move up the gradient and eventually reach the area of highest concentration. This is in fact a relatively
advanced kind of orientation called a *taxis*, in the classification of Fraenkel and Gunn (chemo-

taxis isn’t a real taxis in their terms\(^7^4\)).

Unfortunately for bacteria, simultaneous comparison is limited by the physical separation of

the two receptors: the closer they are, the smaller the concentration difference that must be
detected. Bacteria are very small, so that their chemical sensors (assuming they have more than

one) must be very close together. Calculation shows that *E. coli* needs to be able to detect con-
centration differences as small as one in 10,000 to move up the gradients in Pfeffer’s petri dishes.

Moreover, if the bacteria could do this, they would be able to go more or less directly up the gra-
dient. But the movements we see don’t look like this: they seem almost random with respect to

the gradient.

The only other possibility is for suc-
cessive comparison. Given some kind of

*short-term memory*, a bacterium might com-
pare gradient concentrations in different

places. If the memory is sufficiently persist-
ent, concentrations at widely separated

places can be compared and quite gradual

gradients detected. Such a process of suc-
cessive comparison could modulate a largely

undirected pattern of move-

ment as concen-

tration increased and decreased. This kind

of orientation mechanism is termed a *kinesis*.

In the 1970s, McNab and Koshland

did several ingenious experiments to show that successive comparison is the key. They

used both *Salmonella* and *E. coli*. First, they

found that the organisms showed the same

random-

walk-

like behavior at different abso-

lute concentrations of an at-

tractant – provid-

ing they were given some time to settle

down (see Figure 4.3). This shows that

concentration by itself has no effect on their

behavior. McNab and Koshland then subjected the bacteria to a sudden, “step” change in con-
centration (with no gradient), by adding new attractant to the water and mixing it rapidly. They

looked at the bacteria immediately after mixing was complete. If the bacterium’s behavior is

guided by simultaneous comparison between the concentrations at its “head” and “tail,” there

should be no effect of the sudden change, because there is no gradient. But if its pattern of

movement is determined by a difference across time, then a sudden increase in attractant concen-

tration should produce the same effect as swimming up a gradient.

The experiment showed that the bacteria use successive comparison and it works like this. Both species of bacteria show only two modes of movement: straight-line swimming, and

“tumbling” – random-walk-like behavior (see inset, Figure 4.2). The organisms shifted imme-

diately from tumbling to straight swimming when attractant concentration was increased. In an-

other experiment, tumbling increased when attractant concentration was reduced. The experi-

\(^7^4\) Terminology: In biomedicine the term taxis, as in *phototaxis*, is often used interchangeably for kineses and taxes. The term *tropism*, coined by proto-behaviorist Jacques Loeb (see Box 4.1) also refers to mechanisms of this type. For consistency, we will stick to the traditional distinction first proposal by Fraenkel and Gunn (1940) based on the earlier, German work of Kuhn. According to this scheme, *taxes* refer to directed orientation, where the animal goes more or less straight towards or away from a light or a sound, whereas *kineses* refer to random-walk-like behavior that is modulated by a concentration gradient.
ments proved that the bacteria detect a concentration gradient by making successive, rather than simultaneous, comparisons.

Gradient-detection is, of course, the answer to the puzzle of M. xanthus predation with which I began: M. xanthus secretes an attractant that draws E. coli cells to them — to their destruction. The sluggishness of xanthus is an advantage, not a flaw, because it allows the unwitting E. coli to find its nemesis. Shi and Zusman entitled their note “Fatal Attraction”.

Short-Term Memory

Over how much distance can bacteria compare? This is determined by a process of sensory adaptation that is in effect a short-term memory. Figure 4.4 shows the process. The graph shows the proportion of Salmonella swimming straight at different times after a step increase in attractant concentration. Right after the change the bacteria “think” they are going up a gradient, so they swim in straight lines (“Things are getting better, so we must be doing right!”). But as time elapses, they adapt to the change, things are no longer improving and the curve declines as fewer and fewer bacteria swim straight. Adaptation is complete after about a minute, when all of the bacteria are tumbling again.

The adaptation mechanism allows these organisms to make comparisons across a distance of between 20 and 100 body lengths, which reduces the difference-detection problem from 10,000 to 1 to between 100 and 1000 to one: still difficult, but much better than simultaneous comparison. The bacteria have reduced a complex problem in three-dimensional orientation to simple on-off control of a random pattern: when things are getting worse, tumble; when they begin to improve, swim straight: “By taking giant steps in the right direction and small steps in the wrong direction, [Salmonella] biases its walk very effectively in the direction which aids its survival” as Koshland puts it (1977, p. 1057) The variation-selection process is particularly clear here: tumbling varies direction of movement; an increase in attractant concentration signals, and selects, the right direction to move.

Box 4.1: Micropsychology: Jacques Loeb (1859-1924) and H. S. Jennings (1868-1947)

At the beginning of the twentieth century, there was great interest in the behavior of microorganisms as a guide to principles that might be applied to the higher animals and man. The two principal players were Jacques Loeb, a Jewish-intellectual émigré from Germany, and H. S. Jennings, a U.S. native, born in middle-class circumstances in rural Illinois.

Loeb’s parents died while he was young but left him with resources sufficient to go to university (then open only to a privileged few) and enter on an academic career. He studied medicine and did research on brain function in dogs in Germany. In 1885 Loeb got an assistantship in Berlin but gave it up after a year, mainly because of his growing reluctance to inflict brain lesions on his dogs. He moved to Würzburg and there became interested in the reactions of simple animals, and plants, to light. Out of this rather limited experimental experience developed Loeb’s theory of tropisms. The ideas were considerably elaborated in later years, but the basic notion was simple: movements such as phototropism are automatic reactions to stimulation. In 1891 Loeb brought these ideas, and himself, to the U.S., when job scarcity and anti-Semitism in Germany made the U.S. an attractive option. He soon got a job at the then-new University of Chicago, where he remained for several years. He later worked at Berkeley and finally went to the Rockefeller Institute (now University).
Herbert Spencer Jennings, son of an evolution-minded physician father (his other son was first-named “Darwin”), went to the University of Michigan where he got a degree in biology. He entered Harvard in 1894 and got a Ph.D. in two years for work on the early development of a rotifer, a multicellular microscopic animal. He did research at Jena (home of the famous Zeiss optical company) in Germany, and Naples, where he did much of the work on the behavior of protozoa later written up in his classic work *The Behavior of the Lower Organisms* (1906). Jennings returned from Europe to become professor of botany at Montana State A & M, then instructor in Zoology at Dartmouth and then instructor and assistant professor at the University of Michigan. In 1903 he went to the University of Pennsylvania as assistant professor of zoology. In 1906 he went to Johns Hopkins as associate- and then full-professor of experimental zoology. He remained at Hopkins until his retirement in 1938. After finishing his book on behavior, Jennings become more and more interested in genetics.

Loeb had a greater influence on the early development of psychology than Jennings, probably because his ideas were simpler, and he readily provided prescriptions for action. His tropism theory was based on two simplifying assumptions that Jennings, quite correctly, resisted: that an organism is passive, remaining quiet unless forced into action by an external stimulus; and that the reaction to a stimulus does not depend on the state of the organism. Loeb’s conviction that the behavior of microorganisms is exceedingly simple and automatic led naturally to preoccupation with the control of behavior. He was not sympathetic to placing animal psychology within an evolutionary context. John B. Watson, passionate promoter of behaviorism, later found Loeb’s ideas an easy model for his own highly simplified and fiercely deterministic view of behavior.

In contrast to Loeb, Jennings’ ideas were based on extensive and careful observations and experiments on protozoa. He noticed that spontaneous activity far from being a rarity is the rule — it is reflex-like behavior that is the exception — so that Loeb’s idea of “forced action” could not be generally correct. We now know that variability in behavior is essential to adaptation to unpredictable conditions: a “Loebian” animal could learn nothing new. Like Loeb, Jennings saw many similarities between the behavior of protozoa and higher animals, but unlike Loeb he drew from this the conclusion that both were complex, not that both were simple: “[I]f Amoeba were a large animal, so as to come within the everyday experience of human beings, its behavior would at once call forth the attribution to it of states of pleasure and pain, of hunger, desire and the like, on precisely the same basis as we attribute these things to the dog.” Nevertheless, contrary to the assertions of famous historian of psychology Edwin G. Boring in his influential text (1929/1957), Jennings was not arguing from this that protozoa were “conscious”, only that simple stimulus-response rules (although this terminology is from a later time) are not adequate to describe their behavior. His idea that much behavior is spontaneous, in the sense that it is not caused by a current stimulus, is an acknowledged anticipation of B. F. Skinner’s idea of emitted operant behavior. But Jennings was an experimenter rather than a theorist: he could see what was wrong with Loeb’s tropisms, but he lacked the theoretical tools necessary to provide a testable alternative. We now possess the tools: The computer makes it relatively easy to reveal the often-complex behavior of simple processes; and modern molecular biology offers techniques for dissecting their physical basis (see, for example, Pennisi, 1999) — a real possibility here, if not in the case of learning in mammals and birds. It remains to be seen whether Jennings’ imaginative program for the experimental analysis of protistan behavior can now be brought to fruition.


As we move from bacteria to more complex single-celled animals, behavior becomes correspondingly more complex. H. S. Jennings (see Box 4.1) described the behavior of ciliates such as *Paramecium* and *Oxytricha*, much larger and more differentiated cells than bacteria. He observed how these animals avoid a region that is too hot, too acid or too poisonous. For example, when *Oxytricha* first enters an unfavorable region the animal stops, backs, then turns and starts to move in a new direction; if the bad area is encountered once again, the organism repeats the process until it finds a successful direction. The direction of the turn each time is unrelated to the actual direction of the boundary. This trial-and-error mechanism is obviously similar in many ways to the initiation of tumbling by bacteria when circumstances change for the worse. It is a bit more efficient, because the first reaction of the animal is approximately to turn away from the boundary. The process is far from perfect, though, and the animal may spin around several times if it encounters something really bad – showing that the reaction is automatic: the worse the change the more spinning. The process would be more efficient if *Oxytricha* could turn reliably away from the boundary, but this would require that it know where on its body the bad stimulus is, which is presumably beyond its sensory capacity.
Paramecium and many other ciliates show essentially the same pattern as Oxytricha, reacting to negative gradients by retreating and turning. As with the bacteria, this process serves both to avoid bad regions and seek good ones. If a bad region surrounds a good one, the process keeps the animal within the good region.

THE DYNAMICS OF TRIAL-AND-ERROR
Chapter 1 introduced the topic of dynamic models. Now I take it up more formally, beginning with the basic concepts of dynamics and then going on to the simple ingredients needed for a real-time dynamic model of kinesis and other trial-and-error behaviors.

Model Types
As I noted in Chapter 1, the word “model,” like the word “function,” is used in many senses in science. I use it here to mean a well-defined (i.e., not merely verbal) theory. The differences between a model and a theory are comprehensiveness, fixity and expectation of permanence. We say “quantum theory,” because the theory is comprehensive, fixed in form and unlikely to be superseded. I will say “habituation model” because its scope is limited, its details still subject to modification and its future uncertain. But theory and model refer to the same sort of thing and I will use the terms interchangeably.

Models and theories can be either static or dynamic. A static model lacks any reference to time or process. The economic theory of supply and demand discussed in Chapter 3 is a typical static model. The theory simply asserts that for a population of consumers and producers there exist two fixed functions: a demand function and a supply function, and that the equilibrium price and amount supplied will be at the intersection of the two functions. The principles discussed in Chapter 3 are part of comparative statics, the area of microeconomics that deals with equilibria based on individual preference. Statics is so called because it says nothing about how or when the equilibrium is achieved (a dangerous omission, as we will see).

Dynamic models, which I also call mechanisms, deal with process: how a given behavior comes about and its time course. There are three types of dynamic model in psychology.

- **Trial-level** models define a process that operates trial-by-trial, that is, the variables of the model are updated on each trial. Trial timing is determined by the experimenter and is not part of the model. The well-known Rescorla-Wagner (1972) model of Pavlovian conditioning is a model of this type. It describes how the “associative strength” of different stimuli changes from one learning trial to the next as a function of whether reward or nonreward occurred on a given trial.

- **Response-level** models describe how the model variables change from response to response, depending on events such as reward and nonreward (the timing of responses is determined by the experimental subject and is outside the model).

- **Real-time** models describe how the model variables change in real time. Real-time models are obviously the most desirable, because they leave nothing out.

The equations for a dynamic model can be either continuous (i.e., differential equations) or discrete (discrete-time equations). This difference has more to do with computational convenience than model type. Discrete-time equations are handy because they can be programmed directly on to a computer; but there are also standard techniques for discretizing differential equations, so the difference is not crucial.

Dynamic models can be either truly dynamic or quasi-dynamic. Quasi-dynamic models describe only the average direction of change in the variables of the model, $\Delta y = f(x)$, where $f(x)$, is known imperfectly. Quasi-dynamic models may permit the derivation of equilibria, but do not allow the prediction of moment-by-moment changes in behavior (Staddon, 1988a). Truly dynam-

---

75 Several of these distinctions were first made in an important PhD dissertation by Richard Sutton (1984).
ic models predict behavior iteration by iteration (e.g., trial by trial, response by response, time step by time step).

**Equilibria**

Any dynamic system left to run for a long period of time without external intervention will settle down to some pattern of behavior, termed the *steady-state* or *equilibrium* pattern. For example, a real pendulum (not the frictionless pendulum discussed in Chapter 1) will eventually come to rest after a displacement. This single value is called a *fixed-point* equilibrium. Fixed-point equilibria can be *stable*, *unstable* or *neutral*, as defined by the system’s reaction to small perturbations. These distinctions are conventionally illustrated by the image of a ball in a cup (stable), on a cup (unstable), or on a flat surface (neutral), as in Figure 4.5. A stable equilibrium is restored after a perturbation, an unstable equilibrium is lost and a neutral equilibrium shifts by an amount related to the size of the perturbation.

Equilibria need not be fixed points. The system variables can trace out a closed figure (e.g., line, circle), termed a *limit cycle* or *attractor*, or a figure that is determined but not periodic. A nonperiodic attractor like this is termed *chaotic* and this type of behavior is called deterministic chaos. I give examples of a fixed-point equilibrium and limit cycle in a moment. Chaotic attractors are discussed later. Multidimensional systems can show mixed forms, such as *saddle points*, which are stable equilibria with respect to perturbations of one variable, unstable with respect to one or more others. (See Abraham & Shaw, 1992, for visual illustrations of dynamical behavior, including many different types of approach to equilibrium.)

**Simulation**

Anyone who has learned some math knows that a mathematical proposition looks very different after you understand it than it did before. What afterwards seems obvious, may before have appeared completely incomprehensible. How to get from before to after? There is no substitute for practice, for actually working some examples. But in psychology, mathematics is supposed to be incidental to the “real” stuff, the facts and suitably labeled processes of learning, memory, attention or motivation, or whatever — so courses and textbooks rarely devote much time to it. Big mistake — because the real stuff rarely bears close scrutiny and turns from stuff to fluff when we try to model it formally. Most purely verbal explanations for dynamic phenomena in psychology turn out to be more or less unsatisfactory (e.g., Staddon, 1984b).

Fortunately, the mathematics of the phenomena I have been talking about are very simple and contemporary computer software makes it easy to simulate the kinds of processes I discuss. Working with a simple program can usually be as effective as lengthy formal analysis, if the objective is to understand what is going on, rather than to achieve skill at algebra. Moreover, most

---

76 This need for practice in understanding mathematical arguments is very odd. After all, the process to be understood is supposedly just a matter of logic: calculus or limits or Newton’s method for extracting square roots involve nothing but algorithms, well-defined rules for achieving well-defined results. Yet most students cannot grasp mathematical processes like these until they have worked a few examples. The first few examples may be done with little comprehension, perhaps. But the act of doing them somehow induces comprehension later. One conclusion (which will not be congenial to some varieties of cognitive psychology) is that we do not learn by reason alone. Another conclusion (which will not be congenial to naïve behaviorism), is that habits can by some means give rise to ideas and understanding.
of the processes I discuss are nonlinear: they involve if-then thresholds, or multipliers or higher powers. Nonlinear processes cannot usually be solved analytically. Simulation is the only way to understand them. But a nonlinear process may nevertheless be conceptually simple and therefore easy to understand.

Simulation can be done using a standard programming language, like C++ or PASCAL, or using one of the several packages designed explicitly for mathematical analysis, such as MATHCAD or MATLAB. But the simplest method is to use what has become a universally understood programming language: the EXCEL-type spreadsheet. All spreadsheets have graphing as part of the package, so that the results of computations can be seen immediately, often right next to the formulas. The simulations I discuss next were all done this way. I describe four: the discrete-time pendulum, the leaky integrator, tumble-and-swim orientation and Stentor escape behavior.

**Discrete-time pendulum**

The dynamics of a pendulum, so difficult to capture in words, can be expressed in a trivial discrete-time equation:

\[
x(t+1) = x(t) - x(t-1),
\]

where \( x(t) \) is the displacement at time \( t \). The equation simply says that the position, \( x \), in the next time step is equal to the difference between the positions in the previous two time steps. It’s not obvious that this equation yield pendulum-like behavior, but as we’ll see, it does.

We need to pick values for \( x \) at time zero and time zero plus 1. They are the initial conditions of the system. With initial conditions of \( x(0) = 1 \) and \( x(1) = 1 \), the system behaves as in Figure 4.6. The figure shows two graphs. The upper graph shows the behavior of a discrete-time version of a frictionless pendulum as a function of time step. From one time step to the next, the single variable of the system, \( x \), the displacement of the pendulum from its central, resting position, moves back and forth along the line from -1 through 0 to +1.

In general, the displacement of the pendulum follows a repeating pattern, D,D,0,-D,-D, where D is the initial displacement. The lower graph represents this limit cycle in what is termed a state space. The location of the system in its state space completely defines the system, in the sense that the future behavior of the system (e.g., response to additional perturbations, in this example) will be the same no matter how it got to that particular state-space location. Notice that although there is only one output (response) variable in this system, the state space has two dimensions, corresponding to \( x(t) \) and \( x(t-1) \). This is because the next position of the pendulum depends not just on the current position but the position before that: given \( x(t) = 0 \), the next position is either 1 or -1 depending on whether \( x(t-1) \) was -1 or 1, for example. The integrator, discussed next, has a state space with only one dimension, because the next value of its single variable depends only on the current value. Periodic behavior requires a state space with at least two dimensions.
Time delays make a great difference to system behavior. $x = 0$ is the static equilibrium of the pendulum system, but because of the influence of the previously displacement on the change in displacement, the system only stabilizes at that point under very special conditions: namely, when it begins there and is not displaced. **Figure 4.7** shows what happens when the time delay is increased: unless the damping factor, $a$, is less than .6, the system becomes unstable.

**Short-Term Memory: The Leaky Integrator**

Recall our conclusion that bacterial movement up an attractant gradient requires some kind of short-term memory to allow the moving animal to compare concentration levels at different times (and thus of widely separated places). What kind of memory might that be? Figure 4.4 provides some clues. It shows a smooth decline with time in the number of bacteria swimming straight following a step increase in attractant concentration. Evidently, the memory is not of the on-off computer variety. It seems to be more like a capacitor charging (when the attractant concentration changes) and then discharging after the change. One problem is that a capacitor discharges in an exponential way, whereas the curve in Figure 4.4 is sigmoidal: declining slowly at first, but then more rapidly and more or less exponentially. The “real” memory curve may be exponential nevertheless, because the curve in the figure is an average of the behavior of hundreds of bacteria. Each bacterium is in fact in one of two states, either swimming straight or tumbling, and the bacteria may also differ quantitatively, if not qualitatively, in how they react to attractant. So this is not a case where we can extract the underlying process directly from the data (such cases are in fact rare).

The process of discovery actually involves five steps:

1. Devise a model that relates changes in attractant concentration over time to the two possible behaviors of an individual bacterium, tumble vs. straight swim.
2. Derive predictions of the behavior of the individual bacterium in response to different temporal and spatial patterns of attractant.
3. Make an assumption, or assumptions, about the ways that individual bacteria differ, in terms of the model properties and parameters.
4. Derive predictions for each of these assumptions about the behavior of a group of bacteria: proportion swimming vs. tumbling as a function of attractant distribution.\(^{77}\)

\(^{77}\)“Why not just study individual bacteria directly?” you might reasonably ask. The answer seems to be that this is technically difficult and also impractical (probably) because the behavior of individuals is highly variable.
5. Compare these predictions with group data.

In what follows, I will concentrate on steps 1 and 2. Rather little collective effort has been devoted to steps 3 and 4 in any context. Usually, theorists will leap directly from step 2 to step 5, and in this case, so shall I.

**A Memory Element: The Integrator**

*Figure 4.8* shows the simplest possible analogue memory element, a leaky integrator, together with its response to a pulse input. The inset shows a bucket with a hole in the bottom partly immersed in water. The depth of the water is the dependent variable (response) $V$; the distance between the bucket bottom and the surface of the water outside is the independent variable (stimulus) $X$. The graph shows how this system behaves when the bucket is suddenly pushed below the water surface, held there for a while and then lifted back to the surface. When it’s first pushed down, water rushes in and $V$ increases, at a rate determined by the size of the hole in relation to the size of the bucket. When it’s lifted up again, water rushes out and $V$ decreases at the same rate. The equation for the one-parameter integrator (I’ll omit “leaky” in the future) is

$$V(t+1) = aV(t) + (1-a)X(t),$$

(4.2)

where $a$ is a parameter that determines how slowly the integrator “forgets” – the larger the value of $a$, the more slowly $V$ declines with time when the input, $X$, goes to zero.

It’s easy to see from the physical model that the integrator has a fixed-point equilibrium in response to a steady input: $V(\infty) = X$. (Formally, just set $V(t+1) = V(t)$ in Equation 4.2, and solve.) In a two-parameter version of the integrator (convenient for some purposes, as we’ll see in a moment), the term $(1-a)$ is replaced by $b$. The steady-state value for the 2-parameter version is $V(\infty) = bX/(1-a)$.

The Differentiator. *Figure 4.9* shows what happens when the integrator output, $V$, is subtracted from its input, $X$. The system behaves as a differentiator. The heavy line in this graph ($X-V$) is labeled “*affect?*” because the well-known Solomon-Corbit (1974), “opponent-process” theory of motivation says that emotion works this way: a large effect following stimulus (e.g., drug) onset, that diminishes later, and a negative rebound when the stimulus is withdrawn. This model shows that opponent processes are not necessary to explain these two effects: a single differentiator will suffice. So-called “edge effects” in perception are also examples of differentiation, in this case spatial as well as temporal differentiation. Many motivational and stimulus effects follow this pattern.

The diagram at the bottom adds a threshold to the output of the differentiator, meaning that the output of the system as a whole, $V_o$, is zero unless $X-V$ exceeds some threshold value. I will use this feature to translate the output of this system into either a graded response, in the case of habituation, which is discussed in the next chapter, or the on-off tumble-vs.-swim decision by bacteria, discussed next. This little unit, integrator with summer ($X-V$) and threshold, is the elementary ingredient of many adaptive processes.

**Tumble-and-Swim Orientation**

*Figure 4.10* shows the simulated path of an orienting bacterium in a concentric gradient with a maximum at the intersection of the two axes. At first the organism charges off at a tan-
Adaptive Dynamics

Staddon

59

gent, but after a couple of direction changes, it finds itself in the vicinity of the peak concentration, and remains there, moving in a random walk. At the top of the figure is a diagram of the orientation process. The differentiator-threshold unit is the one just described (Figure 4.9). The threshold ($\theta$, a measure of the sensitivity of the sensory system: the smaller the threshold, the more sensitive) for this simulation is zero, so in effect the system just tumble if the differentiator output is negative (things are getting worse) and swims straight if the output is positive (things are getting better).

Formally: if $V$ (Equation 4.2) $> X$ (current gradient value), then take one step in a random direction; otherwise, take on step in the same direction as the previous step.

This simple system is surprisingly efficient at moving the organism up a gradient. Like any adaptive process, it has its limitations, however. It does not do well at detecting a sharp gradient, for example. Because the “memory” of the system (determined by parameter $a$) is relatively short (for efficiency in moving up smooth gradients) the system may be confused by a sharp, “step” gradient. It will sit as happily on the low side as on the high side, thus losing about half the time.

Avoidance Behavior

Stentor is a highly differentiated single-celled animal that looks like a trumpet (Figure 4.11). Jennings has provided a fascinating detailed account of how this incredibly primitive animal nevertheless manages to try out different options in an effort to avoid a noxious substance in its vicinity:

Let us now examine the behavior under conditions which are harmless when acting for a short time, but which, when continued, do interfere with the normal functions... by bringing a large quantity of fine particles, such as India ink or carmine, by means of a capillary pipette, into the water currents which are carried to the disk of Stentor.

Under these conditions the normal movements are at first not changed. The particles of carmine are taken into the pouch and into the mouth, whence they pass into the internal protoplasm. If the cloud of particles is very dense, or if it is accompanied by a slight chemical stimulus, as is usually the case with carmine grains, this behavior lasts but a short time; then a definite reaction supervenes. The animal bends to one side...It thus as a rule avoids the cloud of particles, unless the latter is very large. This simple...reaction turns out to be more effective in getting rid of stimuli of all sorts than might be expected. If the first reaction is not successful, it is usually repeated one or more times...

If the repeated turning to one side does not relieve the animal, so that the particles of carmine continue to come in a dense cloud, another reaction is tried. The ciliary movement is suddenly reversed in direction, so that the particles against the disk and in the pouch are thrown off. The water current is driven away from the disk instead of toward it. This lasts but an instant, then the current is continued in the usual way. If the particles continue to come, the reversal is repeated two or three times in rapid succession...If this fails to

---

**Figure 4.10** Top: Tumble-and-swim orientation model. Bottom: Typical path generated by the model (gradient peak is at the intersection of the axes).
relieve the organism, the next reaction – contraction – usually supervenes.

Sometimes the reversal of the current takes place before the turning away described first; but usually the two reactions are tried in the order we have given.

If the Stentor does not get rid of the stimulation in either of the ways just described, it contracts into its tube. In this way it, of course, escapes the stimulation completely, but at the expense of suspending its activity and losing all opportunity to obtain food. The animal usually remains in the tube about half a minute, then extends. When its body has reached about two-thirds original length, the ciliary disk begins to unfold and the cilia to act, causing currents of water to reach the disk, as before.

We have now reached a specially interesting point in the experiment. Suppose that the water currents again bring the carmine grains. The stimulus and all the external conditions are the same as they were at the beginning [my italics]. Will the Stentor behave as it did at the beginning? Will it at first not react, then bend to one side, then reverse the current, then contract, passing anew through the whole series of reactions? Or shall we find that it has become changed by the experiences it has passed through, so that it will now contract again into its tube as soon as stimulated?

We find the latter to be the case. As soon as the carmine again reaches its disk, it at once contracts again. This may be repeated many times, as often as the particles come to the disk, for ten or fifteen minutes. Now the animal after each contraction stays a little longer in the tube than it did at first. Finally it ceases to extend, but contracts repeatedly and violently while still enclosed in its tube. In this way the attachment of its foot to the object on which it is situated is broken, and the animal is free. Now it leaves its tube and swims away. In leaving the tube it may swim forward out of the anterior end of the tube; but if this brings it into the region of the cloud of carmine, it often forces its way backward through the substance of the tube, and thus gains the outside. Here it swims away, to form a new tube elsewhere (Jennings, 1906, pp. 174-5).

The avoidance behavior\(^78\) of Stentor as Jennings describes it is remarkable for a single-celled animal with no nervous system at all. A neural-net model does not seem appropriate. It also seems unlikely that such a creature has “cognitions” and “representations,” so the usual cognitive model is unlikely to be correct. Old-style stimulus-response models are also inadequate, since Jennings emphasizes two successive occasions when the animal’s behavior was different even though all external conditions were apparently the same. How might we explain the apparently intelligent behavior of this minimal creature?

The animal has four levels of response to escape from the carmine, each more costly (but also more likely to be effective) than the preceding one: turning away (A) uses little energy and doesn’t interfere with feeding (E); ciliary reversal (B) uses little energy, but is an interruption of feeding; contracting into the tube and waiting for a while (C) is energetic and seriously interferes with feeding; breaking away (D) is most energetic of all, and means abandoning a known feeding site. These five activities are all mutually exclusive.

We don’t know what causes the animal to shift from one mode of behavior to another. Jennings is typically cautious: “…shall we find that it has become changed by the experiences it has passed through…?”. Well, it seems pretty obvious that some change has taken place in the animal; the question is what, and how can it be mod-

\(^78\) Terminology: This behavior would usually be termed escape rather than avoidance, in the technical lexicon of learning psychology, which is mostly concerned with the behavior of mammals and birds. But avoidance is closer to common usage in the Stentor case, and so I use the term here.
eled simply enough to be plausible for a creature with so few computational resources.

It turns out that *Stentor* avoidance behavior can be modeled using the basic memory element I have been describing. The structure of the model is shown in Figure 4.12. I assume that the noxious stimulus acts as input to each one of a set of parallel 2-parameter integrators. Thus, the strength of the \( i \)th response is given by

\[
V_{i}(t+1) = a_i V_i(t) + b_i X(t),
\]

where \( X \) is the stimulus value and \( a_i \) and \( b_i \) are constants that differ for different activities in a way that I will describe. The response is determined in winner-take-all fashion by the integrator with the highest \( V \) value. The parameter values are chosen so that responses later in the avoidance sequence have larger \( a \) values, so their strength decays more slowly, but higher asymptotes, so under continued stimulation they eventually beat out the fast-acting behaviors, i.e.,

\[
a_i < a_{i+1},
\]

and

\[
b_i/(1-a_i) < b_{i+1}/(1-a_{i+1}).
\]

The beauty of this system is that it easily generates sequences in which succession is not invariable, e.g., \( A \) is sometimes followed by \( B \) and sometimes by \( C \), as Jennings describes. The way this works is shown in Figure 4.13. The top line shows the noxious stimulus, which is present (to the animal) all the time, except when it is retracted into its tube (withdraw: \( C \)). The lines below show the strengths (\( V \)-values) of each activity as a function of time step. The horizontal line is the feeding threshold. The response rule is winner-take-all (WTA) and the result of this competition is shown as the string of letters below. The stimulus is on at the beginning and the model, like the animal, runs through its sequence of avoidance activities, first \( A \), then \( B \) and finally \( C \) as the \( V \)-values for all rise at different rates to different heights. \( C \) (retraction) removes the stimulus, and feeding resumes, which restores the stimulus. But then, although “The stimulus and all the external conditions are the same as they were at the beginning”, the model (like the animal) retracts again, rather than again running through the \( ABC \ldots \) sequence. As in Jennings description this is “repeated many times.” Indeed, it will be repeated indefinitely, because I omitted the fourth avoidance response, breaking away, from the simulation so as not to clutter the figure. But a fourth \( V \)-value, slower than \( V_C \) with a higher asymptote, could easily be added, and after sufficient experience with the stimulus it will win and the model, like the animal, will “break away.”

This model has not been tested (*Stentor* research is notoriously under-funded), but it suggests numerous tests and is simple enough to be embodied in the biochemistry of a single cell. The *Stentor* model introduces a theme I will explore extensively in later chapters, namely that the dynamics of a wide range of adaptive behaviors can be explained by the interaction of memory elements with different time scales.
CONCLUSION
Trial-and-error behavior is the prototype for ontogenetic adaptation: adaptive behavior that depends on processes acting during the life of the individual. The ingredients are the same as those for phylogenetic evolution, namely some source of variability, appropriate to the circumstances, a selector, and a process that retains a record of what has been selected, the ontogenetic equivalent of genetic inheritance. For simple organisms, variability is often more or less random, like bacterial tumbling or the twining of plants, but it is often highly constrained, like the avoidance behavior of *Stentor*. Selection usually involves some kind of differentiator, to detect changes for the better or for the worse, although, again the *Stentor* system has nothing that detects change *per se*. It works by emitting a series of more and more costly behaviors that persist until the noxious stimulus ceases. “Good” for this system is just the absence of anything that activates avoidance behavior. And finally, “inheritance” for these simple creatures may be external—the twining plant doesn’t have to remember a good position, it’s *in* a good position and will remain so unless the environment changes. But “memory” is more usually internal: the integrator or integrators whose state captures the recent history of good and bad events. For the tumble-and-swim bacterium, the state of its single integrator tells it how things were in the recent past. Comparing it with the current state of things allows the organism either to persist in its course (if things are getting better) or try another course at random if they are getting worse.

But there is a type of adaptive behavior that doesn’t depend on ontogenetic trial and error, namely *instinctive* behavior, reactions that occur in more or less ballistic fashion in highly specific circumstances. Examples are aspects of courtship and mating behavior, some kinds of predator avoidance, and reactions to noxious stimuli. Each *Stentor* avoidance behavior is an “instinct” in this sense. The simple learning that it shows is not something different from instinct, it is the result of competition among instincts. Indeed, all “higher” learning may be of this sort. William James pointed out many years ago that human beings do not lack instincts, they just have very many of them. The simplest instincts are *reflexes*, which are the subject of the next chapter.
Chapter 5

REFLEXES

Trial-and-error behavior occurs when things are uncertain, when the organism cannot know the appropriate response. But there are situations that are common in the lives of most animals where the correct response is almost always the same: it is always good to recoil from the fire, to attend to an unexpected sound but to cease attending when it is repeated, to constrict the pupil to bright light, to extend a limb when falling, to turn the eyes to a moving object and so on. Reflexes are the mechanisms that deal with situations like these. They have evolved similarly in all mammals and birds, though whether this reflects inheritance from a common ancestor or ancestors (homology) or common, but independent, adaptations to common situations (analogy, convergent evolution) is not clear.

Reflexes are a rather small part – and not, perhaps, the most interesting – of the adaptive repertoire of higher organisms. Nevertheless, they have historically had a special role because of the apparently simple relation between stimulus and response. Reflex stimulus and reflex response look like cause and effect. Finding causes is the main aim of experimental science, so it was natural to look to reflexes as potential building blocks for “higher” mental functions. The fact that reflexes act rapidly and are relatively independent of the animal’s past history (especially with “spinal” preparations) also makes them easy to study. Moreover, the obvious adaptiveness of reflex properties, and the similarity between reflex properties and some aspects of the properties of higher behaviors, suggested to many early physiologists and psychologists that a complete psychology might be built out of reflexes. This idea now looks naïve and the study of reflexes is a minor part of behavioral psychology. Nevertheless, they have some intrinsic interest, they share many properties with more complex processes, and their dynamics can be simulated using the ingredients I described in the last chapter.

Box 5.1: Historical Note: René Descartes (1596-1650) and the Animal as Machine

Descartes was a French philosopher chiefly famous for the Cartesian “method of doubt” and his stunning discovery of the relationship between algebra and geometry. His skepticism led him to trust only his own thoughts, hence his famous aphorism cogito ergo sum. His insight about the relation between geometry and algebra led to “Cartesian” coordinates. His contribution to psychology reportedly grew out of his acquaintance with the ingenious mechanical devices then common in public parks. Bronze nympha and satyrs, hydraulically activated by a footfall, leapt out in front of startled strollers. Bronze nymphs and satyrs, hydraulically activated by a footfall, leapt out in front of startled strollers. Bronze nymphs and satyrs, hydraulically activated by a footfall, leapt out in front of startled strollers. Crude as these automata seem now, they looked strikingly lifelike at the time (advanced technology is indistinguishable from magic, said Arthur C. Clarke in a memorable comment. He might have added that old technology is about as magical as the magic recipe book). These gadgets apparently gave Descartes the idea that animals are just elaborate mechanical devices. Reflexes apart, humans were excluded, apparently for religious reasons.

Descartes grew up with modest private means that allowed him for the most part to walk at his own pace and think his own thoughts. He was a late riser – he used the mornings for rest and reflection – and loved warmth (“Socrates used to meditate all day in the snow, but Descartes’s mind only worked when he was warm” comments Russell, 1961, p. 543). Anxious to avoid interruptions, even in an age much less intrusive than the present, he moved frequently and kept his address secret from all but a select few. He died relatively young in Sweden, where, out of mistaken respect for royalty, he had taken an uncongenial job as tutor to the energetic, early-rising and cryophilic Queen Christina only a few months before. His great work is the Discourse on Method, which contains his account of analytic geometry; the Treatise on Man contains his contributions to psychology and physiology.

The Reflex Concept: Some History

The reflex idea began in the modern era with philosopher René Descartes (Box 5.1), who was perhaps the first to propose that all animal, and much human, behavior can be explained mechanistically. The older view was that behavior was directed by an unknowable internal enti-
ty ("mind," "self," "free will," the "soul," etc.) opaque to direct study. Though extinct in experimental psychology, this fine old tradition continues in the myths of Freud and other psychoanalysts, theories of multiple personality, the psychology of the "authentic self," etc. The reflex concept was elaborated by the Russian physiologist I. M. Sechenov (1829-1905) who also carried out experimental work of the modern type on spinal reflexes, the built-in protective and integrative responses shown by animals in which the higher brain centers have been severed from the spinal cord (spinal preparations). Work on reflexes was extended and summarized by the British physiologist C. S. Sherrington (1857-1960) in an influential series of lectures at Yale University published in 1906.

Sherrington’s research followed a familiar pattern in science. Rather than attempt to investigate sensory and motor integration as a whole, a confusing and impossible task, he tried instead to reduce it to its simplest elements. He worked with spinal animals because they cannot learn, hence behave in a relatively mechanical, history-independent way that can be easily studied. In his stately late-Victorian prose, Sherrington describes the properties of his experimental preparation as follows:

"Experiment today put[s] within reach of the observer a puppet-animal which conforms largely with Descartes’ assumptions. In the more organized animals of the vertebrate type the shape of the central nerve-organ [i.e., the brain] allows a simple operation to reduce the animals to the Descartes condition. An overlying outgrowth of the central nerve-organ in the head can be removed under anaesthesia, and on the narcosis passing off the animal is found to be a Cartesian puppet: it can execute certain acts but is devoid of mind. Thoughts, feeling, memory, percepts, conations, etc.; of these no evidence is forthcoming or to be elicited. Yet the animal remains a motor mechanism which can be touched into action in certain ways so as to exhibit pieces of its behaviour (Sherrington, 1906/1947, p. xi.)"

"With this reduced, decerebrate animal it is possible to study in full quantitative detail the simplest level of reflex, sensory-motor (stimulus-response) integration, free of the complications introduced by spontaneity and "volition" as well as cognition and other associative processes. Spinal reflexes are automatic but they are also adaptive. They make sense as part of the normal life of the animal:

The movements are not meaningless; they carry each of them an obvious meaning. The scope commonly agrees with some act which the normal animal under like circumstances would do. Thus the cat set upright . . . on a ‘floor’ moving backward under its feet walks, runs or gallops according to the speed given the floorway. Again in the dog a feeble electric current (‘electric flea’) applied by a minute entomological pin set lightly in the hair-bulb layer of the skin of the shoulder brings the hind paw of that side to the place, and with unsheathed claws the foot performs a rhythmic grooming of the hairy coat there. If the point lie forward at the ear, the foot is directed thither, if far back in the loin the foot goes thither, and similarly at any intermediate spot. The list of such purposive movements is impressive. If a foot tread on a thorn that foot is held up from the ground while the other legs limp away. Milk placed in the mouth is swallowed; acid solution is rejected. Let fall, inverted, the reflex cat alights on its feet. The dog shakes its coat dry after immersion in water. A fly settling on the ear is instantly flung off by the ear. Water entering the ear is thrown out by violent shaking of the head. An exhaustive list would be much larger than that given here. . . . But when all is said, if we compare such a list with the range of situations to which the normal dog or cat reacts appropriately, the list is extremely poverty stricken. . . . It contains no social reactions. It evidences hunger by restlessness and brisker knee-jerks; but it fails to recognize food as food: it shows no memory, it cannot be trained or learn: it cannot be taught its name. (Sherrington. 1906/1947, p. xi-xiii.)"

"The deficiencies of the reflex animal are that it can respond only to simple stimuli and has no long-term memory. It can neither learn new things nor recall past experiences. How convenient! The assurance that what is observed now is caused by a stimulus that occurred recently makes the researcher’s life much simpler. The absence of spontaneous movement completes the picture. How different from the intact animal, whose behavior now may reflect expe-"
Adaptive Dynamics

Sherrington was a physiologist and his concern was with the functioning of the nervous system. But his work on reflexes has contributed as much to psychology as to neurophysiology. He defined a reflex as a sensory-motor relation involving at least two neurons between receptor and effector, i.e., at least one synapse, the then-hypothetical, now much-studied point of contact between communicating nerve cells. Nevertheless, not all reflexes are monosynaptic and most of his experiments on reflex function were purely behavioral. Sherrington himself recognized that the reflex is not a specific physiological structure: “A simple reflex is probably a purely abstract conception, because all parts of the nervous system are connected together and no part of it is probably ever capable of reaction without affecting and being affected by various other parts.” (1906/1947, p. 7.) Many of the reflex properties he discovered turn up even in organisms lacking a nervous system. The Sherringtonian reflex is a psychological rather than a physiological concept.

REFLEX PROPERTIES
Sherrington’s concept of the reflex is far from the simple, inflexible, push-button caricature sometimes encountered in introductory textbooks. To be sure, there is always a stimulus and a response. But the ability of the stimulus to produce the response depends not only on the state of many other reflexes but also (in the intact animal) on “higher” brain centers, which retain the effects of an extensive past history. For Sherrington, the function of reflexes was the integration of behavior. As parodied by stimulus-response psychology, unfortunately, the integrative function was largely lost and replaced by the notion of reflex as atom. When we come to look at the processes that underlie reflex properties, we’ll see that there are several kinds of “atom” and they are at a deeper level. Integration emerges from the system as a whole.

Reflex properties can be rather arbitrarily divided into two parts: the properties of individual reflexes, considered in isolation; and the properties of reflex interaction. First, I summarize the properties, then I sketch some of the underlying mechanisms, many of which will already be familiar from Chapter 4.

Individual Reflexes
Threshold. The stimulus for any reflex must be above a certain minimum level if it is to elicit a response. This minimum intensity is called the absolute threshold for the response. For example, a minimum level of tone is necessary to elicit the “orientation” ear movement of a dog; a minimum force and duration of touch is necessary to elicit the scratch reflex, and so on. A level below this threshold will not elicit a response. Thresholds are not fixed. They depend on several factors including the stimulus situation, the animal’s state of “arousal,” its immediate past history, the state of other reflexes and (in intact animals) remote past history and present stimulus context.

The threshold is a statistical concept, because thresholds vary even when conditions are constant. For practical purposes, the absolute threshold is usually taken to be the intensity of stimulus sufficient to elicit the response 50% of the time.

Thresholds are measured experimentally. But the concept of a threshold is also part of most models for adaptive behavior. Theoretical thresholds are not the same thing as experimental ones.

Latency and Strength. The time between the onset of a stimulus and the occurrence of the response is termed reflex latency. Reflex latency is often taken as a measure of the strength of the reflex.
of a particular reflex: the shorter the latency, the stronger the reflex. Latency depends on the same factors as threshold and the two often go together. For example, factors that lower threshold usually shorten latency. Latency also depends on stimulus intensity: a more intense stimulus typically elicits a more vigorous response with a shorter latency. Vigor, latency, after-discharge and probability are often taken as equivalent measures of reflex strength. I discuss in a moment why these measures tend to covary.

Refractory period. Reflexes are either tonic or phasic. Tonic reflexes are maintained by continuous stimulation, as in postural reflexes. Phasic reflexes are transient, as in the startle reflex, or repetitive, as in the scratch reflex. After a phasic reflex response has occurred, the threshold of the reflex is elevated briefly or (if it is elicited) the vigor of the response may be reduced. A refractory period is obviously essential to phasic reflexes, otherwise continued stimulation would produce continual elicitation of the response. The function of the scratch reflex, for example, is to produce repeated back-and-forth limb movements in response to a continuous “tick” stimulus. Time for limb withdrawal (flexion) is provided by the refractory period of the antagonistic extension phase of the reflex. Flexor and extensor control are reciprocally inhibitory, so that activation of one first inhibits and then sensitizes the other in such a way that the reflex stimulus elicits not just a single movement but alternating flexor and extensor activation (“scratching”).

At least two processes seem to be involved in refractory period: inhibition from antagonistic reflexes (see successive induction, below) and habituation.

Temporal summation. The second of two successive stimuli, each below absolute threshold when presented in isolation, may nevertheless elicit a reflex response. For example, a cat will turn toward an unexpected sound with its ears pricked forward, pupils dilated, and muscles tensed ready to see something or act rapidly if necessary. This orientation reflex may be elicited by the second of two subthreshold sounds provided the sounds occur in close succession. Frogs strike at moving prey with their long tongues. The movement is completely ballistic: once begun it is not further guided by prey position. The effective stimulus is a small moving or flickering object (i.e., something “buglike”). If the movement is too small or too slow, the animal may fail to strike. But if two sub-threshold stimuli occur in quick succession, the second may elicit a strike (Ingle, 1973). An effect related to temporal summation is sensitization: the fact that under some conditions, repeated elicitation of a reflex may become easier and easier.

Spatial summation. Somatosensory reflexes (such as the scratch reflex) each have a specific region of the skin within which a stimulus can elicit a response. This region is called the receptive field of the reflex. A single light touch may fail to elicit a scratch-reflex response from a cat, but two touches close together on the skin within the same receptive field may be effective.

Temporal and spatial summation interact. Sherrington gives a nice example. He points out that a single piece of card 6 inches long pressed against the animal’s skin for 5 s is rather a weak stimulus for the scratch reflex: it “evokes a reflex of a few beats, which then dies out”. But the corner of the card drawn lightly along the skin for 5 s “evokes a vigorous reflex that continues and outlasts the application of the stimulus. A successive line is more effective as a stimulus than a simultaneous line of equal length and duration.”

Momentum (after-discharge). The response of most phasic reflexes will generally outlast the eliciting stimulus. The duration as well as the magnitude of the after-discharge generally increases with stimulus intensity.

Habituation. Repeated elicitation of many reflex responses eventually leads to a decrease in vigor and loss of stability (“irregularity and tremor” in Sherrington’s preparation) and finally to no response at all (extinction). Habituation to a weak stimulus occurs sooner than to a strong one (a strong stimulus may lead to sensitization). A reflex habituated to a weak stimulus can be elicited again by a stronger one (and sometimes conversely; i.e., some reflexes are highly stimulus-specific). Habituation dissipates with time; this is termed spontaneous recovery.
Reflex Interaction

Many actions physically conflict with one another. You may be able to walk and chew gum at the same time, but you can’t turn left and right simultaneously or scratch with the same limb two different places at once. When actions conflict, the conflict can be resolved either by compromise or domination, by “proportional representation” or “winner-take-all.” Winner-take-all (WTA) is appropriate when an intermediate action is likely to be less “fit,” in the Darwinian sense, than either alternative. For example, if a leopard comes upon two duiker, he had better chase one or the other: heading in between the two is useless. An attacking predator cannot afford to be deflected if a second potential prey intrudes. In many situations, when two incompatible reactions are excited, one must be suppressed if either is to be effective. But there are exceptions. In postural reflexes, for example, a balance between opposing forces may be better than giving in to or one or the other. In assessing the threat potential of a competing male, a stag may do better to weigh a number of factors – his opponent’s size, vigor of movement, loudness and depth of bellow, and so on – rather than relying solely on one. The first two properties of reflex interaction, competition and cooperation, reflect these two adaptive possibilities. The third property, successive induction, has to do with repetitive responses.

Competition. Different reflexes cannot occur at the same time if they involve opposing neural and muscle groups. Even if simultaneity is mechanically possible, the elicitation of reflex A will tend to inhibit reflex B if adaptiveness dictates a winner-take-all type of competition. Thus, incompatible reflexes often show up as behavioral as well as muscular opposites.

Darwin noticed that each emotional “state of mind” tends to be associated with a stereotyped set of reactions – a person who is perplexed may scratch his head, roll his eyes or cough nervously, for instance; whereas an angry person tenses his fists, lowers his eyebrows and bares his teeth. Darwin also pointed out that opposite “mental states” are often associated with physically opposed actions (he called this the principle of antithesis). For example, someone who disagrees with a proposition is likely to shift or close his eyes, whereas if he agrees, he will look straight with open eyes. In greeting his master, a dog tends to show the submissive posture in the upper panel of Figure 5.1: the forelegs are bent, the coat smooth, and the tail is wagging. The dog’s mood when confronting a potential adversary is in a sense opposite to his mood when facing a friend. Darwin pointed out that the active muscles are also opposite: if a flexor is tensed in the first case, the corresponding extensor will be in the second. The bottom panel of Figure 5.1 shows such a threat posture: the dog’s head is erect, his ears pricked, his legs straight, fur and tail raised – all actions opposite to the submissive posture at the top.

Many reflexive expressions of emotion function as communications: to signal “friend” or “foe,” for example. Messages that require very different actions of the receiver – approach vs. withdraw, for example – should be very different in form, to minimize confusion. Thus, recip-

---

80 The same mechanism that allows a predator to detect and concentrate on one prey animal also allows it to be distracted by another. This is one reason prey animals often congregate in flocks or schools.
**Local inhibition** may have a different functional basis in reflexes and in emotional expression. Nevertheless, in both it can be traced to the adaptive importance of decisive action.

**Cooperation.** Reflexes need not always compete, however. Reflexes that share responses may facilitate one another (“allied” reflexes), or when excited at the same time may yield an intermediate response. For example, a subthreshold stimulus for a leg contraction presented at the same time as a stimulus for the scratch reflex may produce leg withdrawal. The same reflex excited by stimuli widely separated within its receptive field behaves in many ways like two separate reflexes. Habituation of the reflex to one stimulus may have little effect on the threshold for the other, for example. Obviously, spatial summation is related to cooperation (and my separation of reflex properties into “individual” and “reflex interaction” is a bit arbitrary!).

Because posture must always be maintained, postural reflexes usually combine synergistically to reach an adaptive compromise.

**Successive induction.** Competing reflexes obviously inhibit one another. But in many cases, cessation of response A may facilitate the occurrence of competing reflex B, and vice versa. This kind of interaction is known as successive induction. It is obviously useful in flexor-extensor interactions in scratching, for example, and in repetitive activities of all kinds, such as walking or flying.

**Figure 5.2** shows an example of successive induction. The response is extension of the hind-limb of a spinal dog in response to a brief electric shock to the paw. Each of the six panels, A-F, shows an elicitation of the cross-extension reflex. The electric shocks are shown at the top. The records below show paw position recorded by a pen attached to the paw writing on a rotating drum: the coordinates of each record are therefore time (X-axis) vs. paw position (Y-axis). The critical comparison is between records A and B. Record A shows a typical modest response (small vertical excursion of the record) in the resting animal. In between Record A and Record B (but not shown) the opposing flexion reflex was strongly excited for 55 s. The next cross-extension response, in Panel B, is much greater than before and doesn’t return to the original level until five minutes or so after four additional elicitations (Panels C - F). Successive induction is an example of an after-effect: elicitation of Reflex A inhibits Reflex B, but offset of A excites B.

**REFLEX MECHANISMS**

“Reflexes” are just a class of stimulus-response reactions with similar properties. Like any classification – the periodic table, in chemistry, Linnaeus’ binary classification of animals and plants, the grouping of soil types in geology – this is just the first part of the scientific story. The next step is to understand why reflexes have the properties they do. The periodic table is explained by
the properties of atoms and electron shells, the mechanism that underlies the rules of chemical combination. Linnaeus’ classification was explained by “modification with descent,” Darwinian evolution. The properties of reflexes are less well defined than the properties of chemical elements, but probably better defined than some taxonomic categories. They do not seem to correspond to simple neural circuits and they occur in organisms with widely different nervous systems. Habituation has similar dynamics in an ant, a worm, a dog and a rat, but the nervous structures involved are not the same. The same goes for sensitization, successive induction, spatial and temporal summation, etc. So, resemblances among reflexes are unlikely to reflect similar neural structures. What is in common among the different organisms is the set of dynamic processes that underlie these various reflex properties. What those processes are is the topic of this section. I’ll begin with habituation and then add to this core to account for the other properties.

**Habituation**

Habituation is the waning of a reflex response, such as “startle” in the rat, to repeated stimulation. Habituation occurs widely in the animal kingdom, from protists to mammals; there are especially good data on worms. The experimental subjects are 1-mm long nematodes, *Caenorhabditis elegans*, a popular species with developmental microbiologists, swimming in a petri dish, observed through a dissecting microscope. Every now and then the dish is tapped. Some of the worms turn in response to the tap. The experimenter measures how many and how much they turn. The average response is plotted as a function of successive taps and a typical result is shown in Figure 5.3. The average response declines with successive taps (stimuli); and the decline is slower and less complete when the taps are widely spaced (60-s apart) than when they are more closely spaced (10-s ISI).

*C. elegans* possesses only 302 neurons and 56 glial cells, and each one has been individually identified. Every cell lineage, from fertilized egg to adult animal, has been perfectly charted. *C. elegans* is not a creature with a great deal of intellectual potential so its behavior dynamics should not be too complicated.

**Figure 5.4** shows one of the simplest possible habituation models, based on a suggestion by Sokolov (1963; Staddon, 1993a). The idea is that the effect of a reflex stimulus in producing a reflex response is the difference between a direct, excitatory effect (the filled-arrow line labeled with a “+” in the figure) and an indirect inhibitory effect (the short gray-arrow line labeled with a “−”) that reflects the organism’s stimulus history (the box labeled “integrator”). Each stimulus (X) adds to the “charge” (V<sub>i</sub>) of the integrator, which therefore builds up at a rate de-

---

---

81 Most of the behavioral data on *C. elegans* and similar preparations refers to group averages. We don’t really know how well a process identified at the group level applies to processes in an individual organism (Staddon, 2019).
determined by stimulus frequency. The net output of the system, \( V_o \), is the above-threshold level of the difference between the stimulus effect, \( X \), and the integrated inhibitory effect of past stimuli, \( V_I \) (combined in a summer, the black dot). This differentiator+threshold process is exactly the same one I described earlier, in Figure 4.9. I have called this a stimulus-type (S-type) habituation model, because the stimulus provides the inhibitory feedback. The other simple type is response-type (R-type), in which the response drives the inhibitory feedback. In fact, for the C. elegans data, the R-type seems to be a slightly better fit, though it is a bit harder to describe (Staddon, 1993; Staddon & Higa, 1996).

**Equations**

Formally, in discrete-time notation, the single-unit feedforward model is described by two equations, a threshold assumption, and the equation for a two-parameter integrator:

\[
V_o(t) = X(t) - V_I(t), \quad \text{if } X(t) - V_I(t) > 0, \quad (5.1)
\]
\[
= 0 \quad \text{otherwise.}
\]

(Equivalently, \( V_O = \theta(X - V_I) \), where \( \theta \) denotes a threshold function).

\[
V_I(t+1) = aV_I(t) + bX(t), \quad 0 < a < 1, \quad b > 0 \quad (5.2)
\]

where \( V_I \) is the integrated inhibitory effect of past stimuli, \( V_o \) is response strength, \( \theta \) is a threshold (zero in all our simulations), \( a \) is a time constant that reflects the period over which past stimuli contribute to habituation, \( X(t) \) is the effect of a stimulus at time \( t \), and \( b \) is the weighting of the stimulus effect. Thus, from one time step to the next, the inhibitory quantity \( V_I \) decreases by an amount \((1-a)V_I\) and also increases by an amount \(bX(t)\), if a stimulus is presented. Note that if \( a \) is small, only recent stimuli contribute, and habituation is rapid if the stimulus spacing is short enough. If \( a \) is large, even stimuli that occurred a while ago contribute: habituation will occur even if stimuli are widely spaced – but habituation is slow. Habituation rate is determined by both parameters, but recovery rate is determined entirely by \( a \) (because \( X \) is zero when no stimuli are presented). Notice also that if parameter \( b \) is large, or the threshold, \( \theta \), is high, response strength may be sufficiently reduced by stimulus presentation that a second stimulus immediately following will fail to elicit the response. Many reflexes show a refractory period of this sort.

Before explaining this model further, I need to define four key terms: response strength, reflex, reflex strength and reflex state:

**Response Strength** is the vigor of the actual response (startle, reversal [nematodes], blink) to a reflex stimulus (a touch, a tap, an air puff), usually measured in physical units, such as stabilimeter force reading, turning amount, or eyeblink excursion. Response strength is often normalized as a proportion of the maximal response.

**Reflex** is the process that intervenes between a stimulus and the response that the stimulus typically produces. A reflex may be physical (i.e., the actual physiological mechanisms involved), or theoretical (i.e., a dynamic model). If the reflex is theoretical, then we can specify exactly its strength and its state. All my statements about internal states, etc., refer to the theoretical reflex, not the actual physiological mechanism.

**Reflex Strength,** \( V_S \), is the potential response strength at any time, given a standard stimulus (set by convention at unity). Reflex strength is the response strength that would have been measured, had a stimulus been given at time \( t \). Thus, when a stimulus is given, response strength equals reflex strength. Note that this definition of reflex strength is strictly empirical and does not depend on any particular model of habituation. Of course, any dynamic model must predict reflex strength, but theoretical values that are below-threshold or negative will have no direct behavioral counterpart. Reflex strength is usually assessed by converging measures: vigor, latency, and probability. I show in a moment why these are related.
Reflex State is a theoretical concept. It is simply the values of all the variables that define the underlying dynamic process. A one-unit habituation model has only a single state variable, $V_i$; $N$-unit models, discussed later, have $N$ state variables. Note that “strength” is necessarily unidimensional, whereas “state” is almost always multidimensional.

When the stimulus presented is the standard, unit stimulus, response strength is equal to the suprathreshold part of reflex strength: reflex strength for the single-unit, zero-threshold model is just $V_o = 1 - V_I$, where $V_I$ is the state of the single integrator.

The way the habituation model works is shown in Figure 5.5. The figure shows reflex strength ($V_o = 1 - V_I$) of a single-unit model, with a unit stimulus presented every second time step for 100 time steps, followed by 100 time steps of spontaneous recovery. Each upward spike of the record is recovery in between stimuli, each downward spike is habituation caused by stimulus presentation. The value at the tip of each above-threshold spike is response strength during each stimulus presentation.

Clearly this model explains the two basic habituation effects: decline in response strength with repeated stimulus presentation, and slower, less-complete decline when stimuli are further apart (although I have not cluttered the graph with a second curve). It also shows (as it must) spontaneous recovery when the stimulus occurs every two time steps until step 100.

Parameters: $a = .9$, $b = .2$.

Figure 5.5 Reflex strength in a one-unit stimulus-type habituation model. A stimulus occurs every two time steps until step 100. The value at the tip of each above-threshold spike is response strength during each stimulus presentation.

Figure 5.6 How the integrator + threshold produces temporal summation. Top: The single peak on the left shows the effect of a brief, isolated stimulus, which is insufficient to drive reflex strength (integrator output) above threshold so elicits no response. But the two closely spaced stimuli on the right do push reflex strength above threshold and produce a response that lasts for three time steps (filled squares). Bottom: Integrator-threshold unit. Integrator equation: $V(t+1) = .9V(t) + .1X(t)$, $X = 0$ or $.1$, threshold, $\theta = .12$.

Figure 5.7 Integrators in cascade. The effect of a single unit spike input on the response of three one-parameter integrators (Equation 4.2) in series. The output of Integrator 1 is the input to Integrator 2, etc. The three curves show the output of Integrator 1 (“First”), Integrator 2 (“Second”) and Integrator 3 (“Third”). The three time constants are $a_1 = .5$, $a_2 = .75$, $a_3 = .95$. Insert: Integrator cascade.
stimulus series is over. (There is a third property of habituation that the one-unit model does not account for, rate sensitivity. I return to rate sensitivity in the next chapter.)

**Temporal Summation, Probability and After-discharge.** Recall that for most reflexes a repeated just-sub-threshold stimulus may elicit the reflex if the separation between the two occurrences is brief enough. Temporal summation is a basic property of any integrator-threshold system. The way that an integrator and a threshold produce temporal summation is illustrated graphically in Figure 5.6, which shows the effect of a single isolated brief (one-time-step) stimulus, on the left, and two closely spaced stimuli, on the right. Both the isolated stimulus and the first stimulus of the pair fail to elicit a response, but the second stimulus of the pair drives reflex strength above threshold and does elicit a response. Notice that the response outlasts the stimulus: after-discharge is also a built-in property of the integrator-threshold system.

Long post-stimulus delays, and up-and-down (rather than just “down”, as in Figure 5.6) changes in response strength imply the existence of cascaded integrators. An example is in Figure 5.7, which shows a series of three integrators (inset), with increasing time constants (legend). The graph shows the response of each integrator in the cascade to a single brief “spike” stimulus as input to the first integrator. The sharp input is progressively delayed and “smeared” through the series: the peak of the final integrator response, for example, is delayed eight time steps after the stimulus. This cascade process by itself shows both temporal summation and (given a threshold) an inverse relation between latency and stimulus intensity. We will have more use for cascaded integrators when I discuss rate sensitivity and feeding regulation in later chapters.

The effect of stimulus intensity on latency is shown in Figure 5.8. A larger stimulus causes a more rapid rise in reflex strength, which therefore crosses the response threshold sooner (i.e., has a shorter latency: arrows). If the input to the habituation-sensitization system is not direct but via an integrator, then the system as a whole will show covariation between stimulus intensity and latency: the stronger the stimulus, the shorter the latency.

“Habit Strength.” In his classic compendium on experimental psychology, C. E. Os- good discussed the relatively weak positive correlations between reflex latency, frequency and amplitude (vigor) as measures of what was then termed habit strength. His comments are interesting for what they say about how psychologists tend to reify empirical terms:

If it turns out that quite different estimates of habit strength can be obtained by using different indices, it becomes reasonable to ask which is the best index of habit strength. Since we have no independent criterion here – no objective, invariant correlate of habit strength – this question cannot be answered with any degree of confidence…” (1953, p. 327)
The point, of course, is not to find the best index of “habit strength,” whatever that may be, but to discover the process that causes latency, frequency and amplitude to covary in just the way that they do in response to stimuli of different strengths and temporal spacings. It turns out that the simple system in Figure 5.6 has the right properties. In this system, there is a single variable, \( V \), that we might term “strength”, but when this system is part of a larger one, many variables are required to specify the state of the whole system and latency, probability and amplitude measures reflect the action of all of them — hence they will be only imperfectly correlated.

If we add as a “front end” to the habituation model in Figure 5.4 the integrator-threshold unit in Figure 5.6, we get a habituation model that shows almost all the dynamic properties of real reflexes: temporal summation, covariation of latency, probability and amplitude; afterdischarge – plus one other property that is normally thought to represent a process different from habituation, namely sensitization. Figure 5.9 shows the responses to a closely spaced series of stimuli (line, filled squares) and a more widely spaced series (gray squares). This system shows sensitization to the closely spaced series (the elevated responding at the beginning of the series) but not to the widely spaced series, a common pattern.

There are several other arrangements of two integrators that show sensitization and habituation. Figure 5.10 shows another example. Integrator 1 with its negative feedforward and threshold is the Sokolov-type habituation model in Figure 5.4. Sensitization is produced by positive (excitatory) stimulus feedforward from Integrator 2, \( V_e \), which is added to the stimulus effect, \( X \), and the (inhibitory) output of Integrator 1, \( V_i \). There is a new type of element in Figure 5.10, the oval circling the output line from Integrator 2, which is a nonlinear element called a gate. Its function is to block the output from Integrator 2, except when
a response occurs. Without the gate, the output from Integrator 2 would allow the response to persist after the stimulus. After-discharge occurs in many reflexes, as I’ve described, but not in all.

Gating can be accomplished in a number of ways: for example, by a logical element (e.g., “if \( X > 0 \), then \( V_e \), otherwise, 0”), or simply by multiplying \( V_e \) by the value of \( X \), so that the combination is zero except when a stimulus is occurring. I have adopted the multiplication method because it allows the model to duplicate other properties of sensitization. The parallel system in Figure 5.10 behaves quite similarly to the serial architecture whose behavior is shown in Figure 5.9.

**Reflex interaction**

I have so far dealt only with individual reflexes and two stimulus properties: frequency (and duration) and intensity. I have not considered the structure of the stimulus or resemblances between stimuli. But when we look at the interactions between reflexes, the stimulus structure can no longer be ignored. The same stimulus may affect more than one reflex and the degree of overlap depends on the way the stimulus is encoded by the sensory apparatus and higher brain structures. I will just discuss the simplest kind of reflex interaction, *dyadic* interaction, the interaction between a pair of reflexes.

Reflexes can cooperate (excitatory interaction) or compete (inhibitory interaction). The interaction can take place between inputs (Stimulus-Stimulus interaction), between outputs (Response-Response interaction), or crossed (S-R and R-S, i.e., between the output of one reflex and the input of the other). There may also be interaction between internal variables, i.e., any of the \( V \) values of one reflex may have an effect on any of the \( V \) values of the other. Three simple possibilities are diagrammed in Figure 5.11. I will look in more detail at two possibilities in **Figure 5.11**: cooperative S-S interaction, and competitive R-S interaction.

**Cooperative S-S interaction.** The idea here is that the stimulus for one reflex also has a direct (but lesser) effect on another. For example, a touch at one point on the skin (which elicits a scratch reflex) may reduce the threshold for (or enhance the magnitude of) a response to a touch at a neighboring point. If the two reflexes affect the same response, this is an example of spatial summation. If we look at the effects of stimuli on a line on the skin on either side of the stimulated point, then the effect on adjacent reflexes will diminish with the distance between the stimulated point and the test point: “In the scratch reflex the mutual reinforcing power between the reflexes falls as the distance between receptors of the arcs increases.” as Sherrington puts it. This graded effect – a facilitation of responding to stimuli that resemble the test stimulus in proportion to their resemblance – was called by Sherrington *immediate spinal induction*, but the modern term for effects like this is *stimulus generalization*. Stimulus generalization (measured in a different way and sometimes termed *stimulus control*) is also an important phenomenon in associative learning. Stimulus generalization is illustrated schematically in **Figure 5.12**.
Competitive R-S interaction. Figure 5.13 shows two gated habituation units (Figure 5.10) connected competitively. The gate (the ellipse around the output from each integrator, near the labels “\(V_{i1}\)” and “\(V_{i2}\)”, on a line connected to the stimulus) functions as a multiplier, so that output from each integrator is zero if its input is also zero. The two habituation units are connected competitively so that the response of one unit is fed back as an inhibitory input to the other. The gates are necessary because without them, elicitation of Reflex 1, say, would cause a negative input to Reflex 2 which, in the absence of a stimulus to Reflex 2, would be subtracted from zero, yielding a positive output from the Reflex 2 summer (black dot) and thus a response from Reflex 2 in the absence of any stimulus. This doesn’t happen with reciprocally inhibitory reflexes.

The equations for Reflex 1 are as follows:

\[
V_{o1} = X_1 - V_{i1} \quad \text{if} \quad X_1(X_1-V_{o2}) > \theta, \quad 0 \quad \text{otherwise.}
\]  

(5.3)

\[
V_{i1}(t+1) = a_1 V_{i1}(t) + b_1(X_1(t)-V_{o2}(t-1)),
\]

(5.4)

and similarly for Reflex 2. \(X_1(X_1-V_{o2})\) is the gate term that ensures that there is no response in the absence of a stimulus.

Figure 5.14 shows some of the behavior of this system. The filled symbols indicate the response of Reflex 1, the open symbols the response of Reflex 2. The stimuli are shown by the bars at the bottom (black: Reflex 1; grey: Reflex 2). The graph shows two experiments. The first one, on the left, is a test for successive induction (see Figure 5.2). Reflex 1 (in its resting state, i.e., after a long period with no stimulus) is stimulated, followed by four elicitations of Reflex 2, followed by another elicitation of Reflex 1. The result: the response strength of Reflex 1 is enhanced by the interpolated excitation of the competitive Reflex 2 – successive induction.

The second experiment begins in the center, after both reflexes have approximately reached their resting states once again. First, Reflex 1 is repeatedly elicited until it habituates completely (the line indicating reflex strength goes below zero). Then Reflex 2 is elicited four times, followed by one stimulus to Reflex 1. The result: a relatively strong response from Reflex 1.
previously habituated Reflex 1. This response is an example of dishabituation, restoration of a habituated response by the presentation of a different stimulus.

In this situation, successive induction and dishabituation are names for two related properties of the same process, R-S competition. Dishabituation can occur for other reasons, but competition is probably the commonest.

I have given only the very simplest examples of reflex interaction, just enough to show how the basic phenomena can be duplicated by simple combinations of integrators, summers, gates and thresholds (and see Staddon, MacPhail & Padilla, 2010 for another example). Identifying the arrangement involved in a particular reflex or reflexes is a matter of careful reverse engineering: proposing a plausible arrangement, looking at its behavior in response to different stimulus patterns, making modifications when data and prediction fail to match, and then repeating the process. When a process is found that withstands all tests, it needs to be scrutinized and simplified until we find the simplest process – fewest state variables, smallest number of free parameters – that will do the job.

**Conclusion: Reflex Mechanisms**

Reflex properties depend on reflex physiology. Yet the properties of reflexes transcend any particular physiology. Habituation, sensitization, refractory period, successive induction, spatial and temporal summation, are found in creatures from nematode to neuroscientist, across a phylogenetic range so wide that there can be few shared properties of the neural and biochemical circuitry. What, then, are these models models of? As I said earlier, the answer is that they are generic models of process, not necessarily or directly models of the underlying physiology. All automobiles follow the same general design plan (bauplan is the useful German word used in comparative anatomy). But the details vary greatly from model to model. Some have front-wheel drive others rear-wheel drive; some have automatics, others are stick shift; some have four cylinders, others eight; some have overhead camshafts others do not; some are left-hand drive, others right; and so on. These dynamic models are of the generic reflex, just as the dictionary definition of automobile is a definition of the generic car. These models capture the kinds of processes and the kinds of computational elements that are involved in all reflexes. I have put these elements together to simulate the behavior of the generic Sherringtonian reflex. Slightly different arrangements may be necessary for specific reflexes.

---

**Box 5.2: John Broadus Watson and Behaviorism**

John Watson (1878-1958) is one of the more turbulent figures in the history of psychology. Born in the South, his religious upbringing failed to take and after a wild adolescence he attended the local college, Furman University, and then after a stint of high-school teaching, went on to do his doctoral work at the University of Chicago, finishing up with a dissertation entitled “Animal education: the psychical development of the white rat”. Watson was attracted to Chicago because of the writings of the professor of philosophy, John Dewey (1859-1952), an influential and not altogether benign figure in the history of American education. But Watson, showing a fine eye for the Emperor’s new clothes, found Dewey’s philosophy obscure on closer acquaintance and struck out on his own. He was more attracted by the hard-headed mechanism of Jacques Loeb (see Box 4.1) and a younger psychologist named James Angell (1869-1949). Under Angell, he did his thesis work on rat learning, beginning a trend that was to dominate U.S. psychology for several decades.

Watson carried on his psychobiological work at Chicago for eight years, and also followed up an interest in what would now be termed ethology: In 1907 he began a series of visits to the Dry Tortugas, off the coast of southern Florida, to study the natural behavior of terns. These papers, although not among Watson’s most famous, are nevertheless among his scientific best.

After his time at Chicago, Watson left for a prestigious professorship at Johns Hopkins University, where he flourished for several years as one of the most popular (and, by all accounts, best-looking) young teachers. Two of Watson’s young associates, Karl Lashley and Curt Richter, went on to become highly influential psychobiologists. Early in his time at Hopkins, Watson became interested in the exciting behavioral physiology been done in Russia by Bekhterev, Pavlov, and their collaborators. His interest in this Russian physiology, and his own work with infants, led Watson to a new philosophical position, first expressed in a journal article and later in a highly influential
book: *Psychology from the Standpoint of a Behaviorist* (1919), a brilliant, though one-sided work. Watson argued not only that objectively measured behavior is the only legitimate datum for psychology (something now generally accepted as *methodological behaviorism*) but also, and wrongly, that all behavior was to be explained in stimulus-response terms: thought became covert speech, measurable by imperceptible movements of the vocal chords, emotion and purpose were all reduced to muscular movements or to the perception of sensations produced by muscular movement. Nevertheless, I would agree with much of the famous summary statement in the 1913 article that announced his version of behaviorism:

> Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness. The behaviorist, in his efforts to get a unitary scheme of animal response, recognizes no dividing line between man and brute. The behavior of man, with all of its refinement and complexity, forms only a part of the behaviorist's total scheme of investigation.

In contrast to the windy academese that characterized much psychology before behaviorism (and not a little afterwards), Watson's new approach was a clarifying vision. Behaviorism dominated experimental psychology for many years and has several descendants: Hullian and neo-Hullian theory; Tolman's purposive behaviorism; and Skinner's radical behaviorism. But its simplism, and reactions against that simplism, continue to impel psychology from one extreme to another in a series of oscillations that persist to this day.

Watson's career at Hopkins ended after World War I because of an incident that would pass almost unnoticed today. Caught in a bad marriage, he became attached to a bright and attractive graduate student, Rosalie Rayner, with whom he had done collaborative research. His wife, at first complaisant, later obtained evidence of an affair and, urged on by relatives, set her lawyers to work. Threatened with dismissal, Watson left academe at the age of 42 and made himself a highly successful career in advertising. He and Rayner were eventually married, and he evidently found lifelong satisfaction in both his new career and his new marriage, during which both he and Rosalie regaled an eager public with firm, marginally harmful advice on the rearing of children. Advertising's (questionable) gain was psychology's loss: for all his oversimplification, Watson's energy and clarity might have matured into a creative psychobiology. Instead, he is remembered more for a youthful polemic than for his many careful experimental and observational studies.

**SUMMARY**

Reflexes are a wonderful topic for scientific study. In Sherrington's spinal preparation, the organism shows no associative memory and is unaffected by events in the distant past, a perfect subject for within-organism experiment. Experimental science is about cause and effect. The reflex, the simplest empirical stimulus-response unit, is the embodiment of cause and effect. Early experimental psychologists well understood the imperative to break down behavior into elements. But, because of Watson's persuasively naïve behaviorism (Box 5.2), they failed to understand the proper role for theory. Consequently, they seized on an experimental entity, the experimentally demonstrable reflex, as *the* elementary unit of behavior. This was a mistake. Elements in science are almost never directly observable\(^82\). In psychology, as everywhere else, real explanatory elements are at the level of theory, not at the level of experimental observation.

Reflexes are to be found in all animals. They share a number of properties (habituation, sensitization, temporal and spatial summation, etc.). I have tried to show how a handful of simple theoretical elements can be connected together to produce a system that mimics the basic reflex properties. In addition to providing a model for reflex dynamics, such a system also explains the covariation among reflex properties of response latency, after-discharge, vigor and probability, as well as the relations between temporal and spatial summation and between successive induction and dishabituation.

---

\(^82\) The only exception I can think of is the *cell*, which was discovered by direct observation by Robert Hooke (*Micrographia*, 1665) when the compound microscope was invented. Genes and even atoms — and synapses — can now be directly observed, but were in fact first discovered through theoretical inference (by Mendel and Dalton — and Sherrington).
Chapter 6

HABITUATION AND MEMORY DYNAMICS

There are rare individuals whose memory mechanisms are defective in an odd way: it is almost impossible for them to forget. The near-perfect memory of mnemonists gives them extraordinary powers – but also extraordinary problems. In a well-known account, the Russian psychologist A. Luria described the mnemonist “S.” who could effortlessly commit to memory pages of figures or words: “Once we were convinced that the capacity of S.’s memory was virtually unlimited . . . we naturally lost interest in trying to ‘measure’ his memory capacity. Instead, we concentrated on precisely the reverse issue: Was it possible for him to forget?” It turned out that forgetting was indeed a problem: “The problem of forgetting . . . became a torment for him.” S. also had great difficulty understanding what he read: “[T]rying to understand a passage, to grasp the information it contains (which other people accomplish by singling out what is most important), became a tortuous procedure for S., a struggle against images that kept rising to the surface in his mind.” (Luria, 1968, p. 34, p. 113)

The problems faced by “S,” who remembered everything, show that normal memory is not mere recording. Normal people not only classify and generalize, they balance remembering and forgetting in a way that “S” found impossible. Neural network theorist Stephen Grossberg calls the problem of balancing forgetting and remembering the “stability-plasticity dilemma.” It is the dilemma faced by the theorist, and by natural selection, in designing an intelligent system that must both change in response to new circumstances, and remain unchanged in respect of features that persist. The minds of men and animals must be selective both in what they learn and in what they forget.

The problem is negative transfer: the fact the what has already been learned can interfere with one’s ability to learn something new. Worse than not knowing, is knowing what ain’t so. This is probably the main reason that adults have more difficulty with new technology than children. It’s not that they are dumber, it’s that they have already learned to do things a different way. Even more dramatic is the difficulty of learning to speak a second language after the age of ten or so. Again, the problem is not an inability to learn, but an inability to unlearn old speech patterns. A perfectly adapted organism would presumably remember what it has learned only for as long as the information is likely to remain valid or useful. Information useful for only a short time — like the phone number of the dry-cleaners — should be forgotten rapidly. But you need to remember your name, the language you speak, and the faces of your friends and relatives.

In a changing environment, the amount of transfer depends on what is learned, and for how long it is retained. If nothing is retained for very long, it doesn’t really matter what is learned, there will be little transfer, positive or negative, from one task to another. The result will be the same if – like “S” – only specific details are learned, nothing is generalized or learned in categories. Even if nothing is forgotten, there will be neither interference nor “savings” because each memory is unique. But in fact some degree of generalization is not only useful, it is essential. Adam, who ate one apple and found it good, did not have to learn all over again to eat another one.

The distinction between what and how-long is the distinction between structural and dynamic properties of memory. Structural properties refer to what the to-be-remembered stimuli are and how they resemble or differ from others – how they are encoded, in cognitive jargon. Dynamic properties relate to the spacing of stimuli in time and how this affects the magnitude and timing of a recall (response) measure. Structural properties have proved very difficult to understand and undoubtedly differ between species and possibly even between individuals. This chapter is mainly about dynamics, how much information is retained for how long.
The basic questions about memory dynamics are: How does the effect of a stimulus change with time, and with repeated presentation? The first question is about forgetting; the second is about habituation. First, forgetting.

**The Forgetting Curve**

Forgetting curves were first measured by German experimental psychologist Hermann Ebbinghaus (1850-1909) in a series of experiments that would today be unpublishable because he worked with only one subject (himself) and used no statistics. Ebbinghaus was interested in how memory fades with time. He used nonsense syllables (indeed, he invented the idea) because he wanted to get rid of all familiar associations. To measure memory loss he learned lists of 13 syllables until he could repeat each list twice without error. He then tested his ability to recall the lists after various delays. He measured strength of recall by savings, that is, the difference between the times it took to learn vs. relearn the list expressed as a percentage of the original time.

There are all sorts of control problems with this procedure – tests at long delays are preceded by more repetitions than earlier tests, for example, later tests may be contaminated by earlier ones, etc. – so that observed performance may not be a pure measure of the effects of delay. Nevertheless, Ebbinghaus’ results have held up in countless more perfect later replications with both human and animal subjects, other stimuli and other dependent measures and industrial-strength statistical tests (e.g., Wixted & Ebbesen, 1997; Rubin & Wenzel, 1996). Ebbinghaus’ work is an early illustration in psychology of the marginal relevance to scientific discovery of inferential statistics and the between-group method. Some of his data are shown in Figure 6.1. The important qualitative property of the forgetting curve is that it declines steeply at first, and more slowly thereafter. This means that it is not exponential; an exponential curve declines at a constant rate.

**Figure 6.1** Retention function from one Ebbinghaus nonsense-syllable experiment. He tested his ability to relearn at periods ranging from 19 minutes to 31 days.

What is the form of the forgetting curve? **Figure 6.2** is a replot of the Ebbinghaus data in log-log coordinates. The points are fit pretty well by a straight line, which shows that a power function,

\[ S = kr^w, \]  

(6.1)

---

83 Would the human-subjects protocol committees of today require that he give himself permission to be a subject in his own experiment? Would they insist on debriefing? Would the experiment perhaps be judged unethical?
provides an adequate description. \( S \) is savings (memory strength) and \( t \) is time since original learning. \( k \) is a parameter representing the initial strength of the memory and \( w \) is a parameter representing the rate of decay. The points are also close to the log form, \( S = k - \ln t \), because the exponent, \( w \), is small and negative. The log-like nature of the retention function seems to underlie important properties of time-interval learning, as we will see in Chapter 14.

**Jost’s Law.** An important consequence of the retention function is illustrated in Figure 6.3, which shows in linear coordinates the data from Figure 6.1, plus the fitted power function. The heavy solid line is the power function for a memory 5000 minutes older and initially (at time zero in the graph) weaker than the light line, which is the function fitted to Ebbinghaus’ data. Although the two functions have the same exponent, they nevertheless cross (arrow). Before the arrow, the new memory (light line) is stronger, after the arrow the older (heavy line) is stronger. This crossover is a common observation first formalized by Jost (1897) as his Second Law: “Given two associations of the same strength, but of different ages, the older falls off less rapidly in a given length of time.” (Hovland, 1951, p. 649). Jost’s law requires non-exponential forgetting, a retention function that falls off rapidly at first and more slowly later, just as Ebbinghaus found (Simon, 1966). I return to Simon and Jost in a moment.

HABITUATION AND MULTIPLE TIME SCALES
Retention functions give a static, molar picture of memory. They are averages (of many trials or subjects) from experiments with a single pattern of stimulus presentation. They are dynamic only in the sense that they show an effect of the passage of time. To get at memory dynamics directly, we need to address the second question I raised earlier: What is the effect of repeated stimulus presentation? Effect on what? one might reasonably ask. Habituation of a reflex requires a response. But there is no (immediate) response in a memory experiment. Nevertheless, a stimulus must have some persistent effect if habituation is to occur – or if the stimulus is to be recalled later. Either way, with a reflex preparation or in a memory experiment, if there is learning, the stimulus must have an effect on the organism’s internal state. The job for any model is to suggest what form that effect takes. The properties of reflex habituation provide an answer.

**Figure 6.4** Single-unit stimulus-type (S-type) habituation model. Integrated effects of past stimuli, \( V_i \), are subtracted from the direct effects, \( X \), and the above-threshold difference, \( X - V_i - \theta \), determines the response strength \( V_o \).
again in Figure 6.4: decline in response strength with repeated stimulus presentation, and slower, less-complete decline when stimuli are further apart. There is a third dynamic property of habituation that the one-unit model does not account for, rate sensitivity. Habituation is more rapid and complete when interstimulus intervals (ISIs) are short than when they are long. Rate sensitivity (Davis, 1970; Staddon, 1993) is the surprising fact that recovery from habituation is also more rapid after short ISIs. Rate sensitivity has been demonstrated at the cellular level (synaptic depression: Byrne, 1982), in the siphon- and gill-withdrawal reflexes of the sea-slug Aplysia (Carew et al., 1972) and in the turning behavior of nematodes (Rankin & Broster, 1992), suggesting that the underlying process may be relatively simple.

The first two properties of habituation, declining response to repeated stimulation and slower decline with less frequent stimulation, can be duplicated by a process in which response strength is the difference between a constant stimulus effect and a “leaky-integrator” short-term stimulus memory (Figure 6.4: \( \theta \) is a threshold, usually 0, \( X \) is the stimulus input, \( V_I \) is the inhibitory integrator “charge” and \( V_o \) is the response output). As successive stimuli are presented, the accumulated charge of the integrator – “memory” for the stimulus – increases. Since response output is the difference between the integrator charge and the constant stimulus input, output response strength decreases as successive stimuli are presented. If the ISI is too large, however, the integrator “discharges” (“forgets”) in between stimulus presentations and the system does not habituate. Thus, the degree of habituation in the model, as in nature, is less at longer ISIs.

The single-unit model habituates less at long ISIs than at short, but post-habituation recovery is just as rapid. A series of at least two cascaded habituation units (the output of the first being the input to the second, cf. Figure 6.5) is necessary to reproduce the rate-sensitive property (Gingrich and Byrne, 1985; Treisman, 1984), and the time-constant of the second (central) unit must be slower than the time constant of the first (peripheral) unit (Staddon, 1993). Staddon and Higa (1996) generalized this system in two ways: to allow for negative feedback from the response (R-type units: Figure 6.6) as well as the stimulus (S-type: Figure 6.4) and to permit cascades of any length (multiple-time-scale habituation).

The equations for the first (peripheral) unit in an S-type integrator cascade are:

\[
V_o(t) = X(t) - V_I(t), \quad \text{if} \quad V_o > \theta \\
= 0 \quad \text{otherwise.} \quad (6.2)
\]

\[
V_I(t+1) = a_I V_I(t) + b_I X(t), \quad 0 < a_I < 1, \quad b_I > 0 \quad (6.3)
\]

where \( V_I \) is the integrated inhibitory effect of past stimuli (I have omitted the \( I \) subscript to avoid clutter), \( V_o \) is response strength, \( \theta \) is a threshold (zero in all our simulations), \( a_I \) is a time constant that reflects the period over which past stimuli contribute to habituation, \( X(t) \) is the effect of a stimulus at time \( t \), and \( b_I \) is the weighting of the stimulus effect. Note that \( V_I \) represents both

Figure 6.5 A cascade of two S-type habituation units.

Figure 6.6 Response-type (R-type) habituation unit. The response is input to the integrator. Integrator output is subtracted from the stimulus input.
the state and the strength of the one-unit system. The single-unit system does not show rate sensitivity, because recovery rate is determined solely by \(a_1\), which is constant and independent of system history. Rate sensitivity requires at least two units in cascade.

In a cascade, the output of the \(j\)th habituation unit, \(V_{oj}\), is the stimulus input to unit \(j+1\), as in Figure 6.5. The equation for the \(j\)th unit in the cascade is

\[
V_j(t+1) = a_j V_j(t) + b_j V_{o,j-1}(t-1),
\]

(6.4)

The final output is just the output of the last integrator in the chain. Any number of units can be cascaded, depending on the number of different time scales implied by data. In a cascade with all-zero thresholds, it is easy to show that the net reflex strength (see Chapter 5 for definitions) at the last integrator (\(N\)) in the cascade is just

\[
V_N = 1 - \sum_{i=1}^{N} V_i.
\]

(6.5)

Notice that the state of an \(N\)-unit model is a vector of \(N\) numbers, the \(V\) values of the \(N\) cascaded integrators — even though strength is still defined by a single number.

Figure 6.7 shows reflex strength as a function of time for a 3-unit S-type cascaded-integrator model. The lower line at the left is reflex strength during and after 50 stimuli at 2 time-step intervals (interstimulus interval, ISI, = 1). The upper line is reflex strength following a stimulus series at 8 time-step intervals (ISI = 7). Stimuli cease after time step 100.

In Figure 6.7, the spikes in each record indicate stimulus presentations.

Recovery is defined as reflex strength in the absence of stimulation. Thus, the curves after the last stimulus presentation (around time step 100) in the graph in Figure 6.7 show the recovery curves for a 3-unit S-type habituation model. Habituation is complete after 3 time steps when the ISI is 1 and after 17 after training at ISI 7. Nevertheless, recovery from habituation is much more rapid after the shorter ISI. This is rate sensitivity.
Parameter Constraints. The behavior in Figure 6.7 shows rate sensitivity because the model has time constants ($a_i$) that increase along the series: smaller (hence faster) on the input side than on the output side: $a_i < a_{i+1}$, for all $i$. This is a reasonable rule because there is informal physiological evidence that peripheral – sensory – processes seem to act faster than central ones. This progression is in fact necessary for rate sensitivity.

The ordered $a_j$ values control the behavior of the integrator cascade in the following way. When stimuli are far apart in time (long ISIs), earlier (faster) units in the cascade discharge more or less completely in between stimulus presentations. Thus, the state of the system (i.e., the vector of $V$ values for the $N$ integrators in the cascade – recall that reflex strength is just the single quantity $1 - \sum V_i$) during the recovery period is dominated by the $V$ values of “slow” integrators late in the chain. Recovery is therefore slow, because these integrators discharge slowly. Conversely, when stimuli are closely spaced (short ISIs), “fast” integrators early in the chain remain charged in between stimulus presentations, consequently block the signal to the slow, later integrators. Recovery is therefore rapid, because the system state is dominated by the high initial $V$ values of fast integrators early in the chain. I’ll explore the consequences of this process for retention functions in a moment.

A cascaded-integrator model can obviously account for the qualitative properties of rate-sensitive habituation. It can also match the quantitative properties, where appropriate data are available (Staddon & Higa, 1996). An extensive dataset has been published in a series of papers by Catherine Rankin and her associates on habituation and recovery in the popular (with devel-

---

**Figure 6.8 Caenorhabditis elegans.**

**Figure 6.9** Data (filled circles) and predictions (lines) for a *C. elegans* habituation experiment of Broster and Rankin (1994). The animals received either 15 10-ISI trials followed by 15 60 s (top) or the reverse (bottom). Notice that habituation during the 10-s series is faster in the 60 → 10 condition than the 10 → 60 condition in both data and model.
opmental biologists) nematode worm *Caenorhabditis elegans* (Figure 6.8). Despite its simplicity, *C. elegans* shows many of the basic phenomena of non-associative learning, including habituation. Catherine Rankin and her collaborators have studied turning by this animal in response to taps on the petri dish in which it is swimming. A dish containing a single *C. elegans* is tapped at fixed inter-stimulus intervals (ISIs). The amount each worm swims backward within 1 s of a tap (termed a reversal) on each trial is recorded, and scored as a proportion of its response to the initial tap. The group response is the average of these individual proportional responses. The recovery function was estimated by presenting stimuli at progressively longer ISIs after the training series had ceased.

Staddon and Higa employed a variety of strategies to reduce the parameter space for N-stage cascade models to manageable proportions (2 parameters). We concluded that a 2-stage R-type (response-type) model gave the best fit. Figure 6.9 shows a typical fit between simulation and data from Broster & Rankin (1994). The left half of the figure shows response strength (for a group of nematodes) during 30 stimulus presentations with different mixes of interstimulus interval (ISI): 15 intervals at 10-s ISI followed by 15 at 60 s or the reverse. The right half of the figure shows recovery following the stimulus series. The data are quite variable, but the 2-stage model does a reasonable job of describing them, and this was true for all the data from two studies (Rankin & Broster, 1992; Broster & Rankin, 1994).

Figure 6.10 shows that a 5-stage model gives a reasonable fit to long-term habituation data from an well-known sea-slug siphon-withdrawal habituation experiment by Carew, Pinsker & Kandel (1972). In this experiment, *Aplysia*

---

84 Usually a brief (1-s) train of six taps constituted the “tap” stimulus.
lysia were given 10 trials per day for 5 days. Habituation occurred both within each day and across days and this pattern is readily duplicated by the multiple-time-scale cascade model.

HABITUATION AND MEMORY

What does the cascaded-integrator habituation model imply for memory? Spontaneous recovery is the change in reflex strength as a function of time in the absence of further stimulation. Response strength recovers with time as memory for the stimulus fades. Spontaneous recovery implies loss of memory strength. In the cascaded-integrator model, memory-trace strength – recallability of a stimulus – is thus the complement of reflex response strength. When all thresholds are zero, from Equation 6.5, it follows that,

\[ V_M = \sum_{j=1}^{N} V_j , \]

where \( V_M \) is memory-trace strength. Equation 6.6 says that memory-trace strength is the sum of several exponentials, since each \( V_j \) value declines exponentially in the absence of stimulation. This is convenient, because Simon (1966) showed many years ago that simple exponential decay is incompatible with Jost’s law, but the sum of two or more exponentials may be consistent with the law. Simon’s argument is illustrated graphically in Figure 6.11. The light line shows a “strong” memory, laid down at \( t = 1 \). The heavy line shows a “weak” memory laid down later, at time \( t = 5 \) (arrow). At \( t’ = 6 \), the strengths of both are equal, but because both traces are exponential, with the same time constant, they decay at the same rate and remain equal for ever. Thus, simple exponential decay is incompatible with Jost’s law.

Figure 6.12 shows that Jost’s law is compatible with the multiple-time-scale (MTS) model. As in Figure 6.11, the figure shows a “strong” memory trace laid down at time \( t = 10 \) and a “weak” trace laid down later at \( t = 26 \). The two traces are equal at \( t’ = 43 \) (arrow), but thereafter, the older trace (light line) remains above the newer (heavy line). The reason the MTS model obeys Jost’s law is that trace decay is not exponential (the sum of several exponentials is not itself exponential\(^{85}\)). Because of the rate-sensitive property, decay at long times occurs at a slower rate than at short times: older memories decay more slowly than newer ones. Hence, given two memories of equal strength, the older one will gain on the newer with lapse of time.

There is now abundant evidence that human and animal retention (forgetting) functions are not exponential (e.g., Staddon, 1984; Hovland, 1951). Rubin and Wenzel (1996) have gathered data from 210 human experiments and fitted them with 105 different of 2-parameter functions (not including the sum of exponentials, unfortunately). The best fitting functions (which

---

\(^{85}\) Simon (1966) pointed out that the sum of two exponentials with appropriate time constants is consistent with Jost’s law. The cascade habituation model is not limited to two stages, however, and the weightings of the \( N \) exponentials are not fixed, but depend on the system history.
fit most of the data very well indeed) all have the general form of the functions in Figure 6.12, declining rapidly at first, but then more slowly.

**Figure 6.13** shows the same data as in Figure 6.12, but in log-log coordinates, like the graph of Ebbinghaus’ data in Figure 6.2. The curves are approximately straight (the curvature at the end of the curve for the older memory is a range effect, because I only used three integrators; up to a dozen may be necessary to accommodate real memory data). Both power and multiple-time-scale (MTS) functions provide an adequate picture of the retention function.

**History Dependence.** But there is a crucial difference between a simple trace model and the MTS model, and it has to do with history dependence. **Figure 6.14** shows a simulated experiment in which a stimulus is presented every 6 time steps for a total of about 200 time steps. The graph shows memory strength during training and after the stimuli presentations cease (corresponding to the recovery period in a habituation experiment). The decaying trace after time step 200 (i.e., after the system has reached steady state) determines forgetting. The question is, How does the form of the trace change as a function of the training history – as a function of whether the training stimuli were closely or widely spaced? **Figure 6.15** shows traces after three different histories: series of stimuli at 2, 3 and 6 time steps. The trace decays more slowly after the 6-time-step history than after the 2-TS history. The range over which this difference exists is small – after a 16-TS history (not shown), decay is again rapid – but this is because of the small number of integrators in the cascade. The point is that in the MTS model, as in real forgetting, amount of forgetting is determined not just by time and degree of learning at the end of training, but also by the spacing of stimuli during training.

**Sensitization and its Opposite.** In Chapter 5, I showed how the addition of a “front end” integrator(s) can produce sensitization in a single unit habituation system (Figure 5.9). A multi-unit cascade may show the opposite of sensitization when the ISI is short. **Figure 6.16** shows an example: when stimulation is continuous, all integrators get charged at first so that total memory strength reaches a maximum (arrow), but when the peripheral units acquire some charge, the later units get less input than they received at the beginning, hence partially discharge, so that total memory strength decreases slightly in the steady state. “Edge” effects like this are occasionally observed.

**Brain Mechanisms**

I argued as forcefully as I could in Chapter 1 that behavioral theories stand on their own feet. Given the richness of physiology and our relative ignorance of the details of brain function, “biological plausibility” is an unreliable guide to behavioral theory. Behavioral theories are valid to the extent that they describe behavioral data accurately and economically. (All that really matters in science, I suspect, is how much can be explained with how little). Nevertheless, some recent real-time physiological data seem to fit remarkably closely the basic assumptions of MTS
habituation theory. Unsought agreement with physiology is good news, even if lack of agree-
ment is inconclusive, so I will describe these data.

The MTS theory of memory dynamics is based on four ideas:

1. That dynamics of retention can be understood through the mechanisms of habituation;
2. that habituation is a process in which responding is inhibited by a leaky integrator system
driven by stimulus input;
3. that habituation units are cascaded; and
4. that the faster units are on the periphery and the slower ones further downstream.

In a *Science* report, Glanz (1998) describes a study reported to the American Physical
Society by S. Williamson and his colleagues that may have identified physiological counterparts
for the last three assumptions. Williamson’s group
used a SQUID (Superconducting Quantum Interfer-
ence Device) to detect tiny magnetic changes. Their
system recorded maps of whole-brain magnetic activi-
ty that could be updated every few milliseconds. In
the simplest experiment, they looked at brain activity
following a single 0.1-s stimulus: “In quick suc-
cession, over less than half a second, about a dozen
patches lighted up like pinball bumpers, starting with
the primary visual cortex [the input to the visual sys-
tem] in the occipital lobe at the back of the brain.” (p.
37) This activation in rapid succession is precisely
what we would expect from a series of cascaded units,
where the SQUID is detecting changes in
\( V_i \), the acti-
vation of each integrator. In a second experiment that
was in effect a two-trial habituation study, subjects
were presented twice with a brief (0.1-s) checkerboard
stimulus. “[T]hey showed the checkerboard twice, with a varying time interval between the dis-
plays, to see whether the first stimulus had left any kind of impression along the way. For very
brief intervals –10ths of a second – only the areas of initial processing in the back of the brain
fired on the second flash, while the others were silent...But as the interval was increased to 10,
20, or even 30 seconds, the downstream areas began firing on the second flash, with a strength
finally approaching that of the initial pop....The data imply, says Williamson, that each site has a
distinct ‘forgetting time,’ ranging from tenths of a second in the primary visual cortex – the first
stage of raw processing – to as long as 30 seconds farther downstream. (p. 37)” Again, this is
precisely the behavior of our cascade of habituation units. Because the initial units have fast
time constants, they block input to the later, slower units so long as the ISI is short enough that
they have not had time to discharge (“forget”) in between stimuli: hence, no response of the
“downstream” units to the second flash at a short ISI. But when the ISI is long, the initial units
have already discharged, allowing the stimulus to pass through to later units, which can therefore
respond. Williamson continues “The memories decayed with the simplicity of a capacitor dis-
charging electricity– exponentially with time – and the later an area’s place in the processing
queue, the longer its memory time was.” Apparently brain “memories,” are leaky integrators and
thus forget exponentially. Exponential recovery seems to be a reliable property of cortical de-
pression (e.g., Abbott, Varela, Sen, & Nelson, 1995), which is presumably what Williamson’s
magnetic changes were measuring.
**Habituation: Associative or non-Associative?**

I have been treating habituation as if it were a reversible, non-associative process – reversible, in the sense that if left for long enough after a stimulus series, the organism returns to its original state. And non-associative in the sense that the process is the same in any context. Neither of these simplifications is really true. Habituation in most species is context-sensitive (Maldonado, Romano & Tomsic, 1997; Whitlow & Wagner, 1984) and the associative component may be more or less permanent. Some animals will remember indefinitely that they have encountered a given stimulus in a given context and first exposure will have persistent effects. But, as Sherrington pointed out (see Chapter 5), essentially every reflex is subject to interactions. Dishabituation, for example is reduced habituation caused by excitation of a competing reflex (see Figures 5.13 & 5.14); dishabituation may also be caused by motivational changes or pharmacological interventions. No behavioral process in an intact animal operates in total isolation. Nevertheless, the dynamics of a process can be studied under constant conditions and if behavior is under those conditions reversible, we may be able to model the process in a relatively simple way, as I have described here for habituation in *C. elegans* and *Aplysia*. Similar behavior has also been observed in rats (e.g., Davis, 1970). Even indefinite persistence of effect is also compatible with this analysis: the last integrator in the cascade may have a time constant of unity. The beauty of dynamics is that it provides a way of understanding behavior that does not require the impossible. Dynamic analysis does not require that we be able to isolate each process experimentally — which is fortunate, because such isolation is usually tough in practice.

**The Partial-Reinforcement Extinction Effect (PREE).** The dynamic process that underlies rate-sensitive habituation may be involved in other learning phenomena. The partial-reinforcement extinction effect (PREE) is observed under many experimental conditions and has two main aspects: (a) Animals trained to respond for infrequent reinforcement stabilize at a performance level that is generally lower than animals trained with more frequent reinforcement. This finding is not paradoxical and fits in with all standard reinforcement theories. But (b), when reinforcement is discontinued (extinction), partially reinforced animals often persist longer in responding than animals that have been reinforced more frequently. This result is paradoxical both because the partially reinforced animals are responding at a lower level at the beginning of extinction, and because classical reinforcement theories assume that more frequent reinforcement builds more response “strength,” hence more persistence.

The PREE is consistent with rate-sensitive habituation in the following sense. The “stimuli” for the system are just the occurrences of reinforcement during training. In a cascaded-integrator system, memory-trace strength is the sum of the $V_i$ values associated with each integrator. This trace strength corresponds to what in cognitive jargon would be called an “expectancy.” The larger this sum (the greater the expectancy) the greater the tendency to respond. “Expectancy” in this sense behaves in a sensible way: it will be higher when reinforcers occur frequently than when they are infrequent, and zero after a long time with no reinforcement ($V_i = 0$ for all $i$). Expectancy will also decay more slowly after partial than continuous reinforcement, providing a basis for the PREE.

To understand the process, let's just look at a 2-stage cascaded-integrator model. Consider how $\Sigma V_i$ (i.e., $V_1 + V_2$) will change with time when reinforcement is no longer delivered. If reinforcements (stimuli) occurred frequently during the preceding training series (a short ISI), then the level of $V_1$ will be high and the level of $V_2$ will be low, for reasons described earlier. The sum, $V_1 + V_2$, will be high. But when reinforcement ceases, $V_1$ will decline rapidly, because $a_1$ is small, hence the sum – and therefore responding – will decline quite rapidly in extinction. The opposite will occur if the training ISI is long: At the end of training $V_1$ will be small and $V_2$ relatively large, although the sum will be less than with a short ISI. But in extinction, $V_2$ declines slowly (because $a_2$ is large), hence the sum also declines slowly and responding is more persistent. This difference is the PREE. There are other ways to apply the rate-sensitivity idea
to PREE – and the effect is not universal (it is not often found in Pavlovian conditioning, for example, and Nevin, 1988, has reported operant experiments that fail to find it). But it is clear that in principle rate sensitivity provides a basis for the effect.

The successive negative-contrast effect (SNCE) is slower responding under low-frequency reinforcement by animals first trained with high-frequency reinforcement (call this group High-Low: HL) compared to animals trained all along with low-frequency reinforcement (Group LL). The effect is usually transient: after sufficient low-frequency training, the performance of both groups eventually converges. The argument of the previous two paragraphs applies here also. For the HL animals, a drop in reinforcement frequency will lead at first to a drop in the sum of $V_1 + V_2$, because $V_2$ will be low and most of the sum will be contributed by fast-decaying $V_1$. With continued experience, the drop in $V_1$ will be compensated by growth in $V_2$. For the LL animals, on the other hand, $V_2$ will already be high because of the initial training. Thus, in a comparison, at the beginning of the second L period, the HL group will at first respond slower than the LL group.

Interval Timing. Chapter 14 describes how memory-trace strength (Equation 6.6) constitutes a kind of clock that animals may use to time intervals.

CONCLUSION

The dynamics of habituation in a wide range of species provide clues to the process of memory-trace decay. Rate-sensitive habituation implies non-exponential memory decay of precisely the right form to accommodate Jost’s law: that old memories gain in strength relative to newer ones. My proposal is that any stimulus, whether it is the stimulus for an overt reflex or a neutral stimulus that has no obvious behavioral effect, is processed by the same kind of rate-sensitive mechanism. This mechanism determines the properties both of spontaneous recovery (for reflexes) and of retention (for stimuli that elicit no measurable response).

The phenomenon of rate sensitivity seems to underlie a broad range of learning and memory effects: habituation, spontaneous recovery, the partial-reinforcement effect, reinforcement-contrast effects and Jost’s memory law. Rate sensitivity implies the existence of processes with multiple time scales. I have just shown that a simple cascaded-integrator system with 2-5 stages can model the dynamics of habituation and spontaneous recovery in a wide range of species. The same model – perhaps with as many as a dozen stages, as Williamson’s data suggest – may apply to human memory.
Chapter 7

**MOTIVATION, I: FEEDING DYNAMICS and HOMEOSTASIS**

An organism is hungry, eats, becomes satiated and then after a while gets hungry again. The change in state associated with this change in behavior is termed *motivational* (see Box 7.1). Motivational changes are recurrent, reversible and *regulatory* – the animal eats just until hunger is satisfied. Motivation is assumed to be a simpler process than learning — and to exist more or less independently of learning. But the fact is that the changes in behavior caused by learning and the changes caused by motivation cannot always be distinguished. Food has motivational as well as learned effects. Understanding motivational effects is essential if we are not to attribute to learning facts that can be explained more simply.

The relations between learning and motivation go in both directions. The changes in behavior as an animal learns include some motivational effects, and the motivating effects of food can be affected by learning. *Taste-aversion learning* is a familiar example of a motivational effect of learning. A rat is given two foods, one novel, the other familiar. Thirty minutes later, the rat is made sick by injection of lithium chloride. Twenty-four hours later, the animal is allowed to choose between the novel and familiar foods. It shows a strong aversion to the novel food (by comparison with a “control” animal given the two foods, but not made sick). And this aversion is essentially permanent, despite the fact that the novel food is completely harmless (Garcia, Ervin & Koelling, 1966; Rozin, 1976; see Logue, 1991, and Mook, 1987, for many other references). This striking example of “one-trial, long-delay learning” has lessons to teach about adaptation – why the rat, a scavenging, oft-poisoned omnivore, should be so sensitive to novelty and such a quick learner even if consequences are delayed; and about how feeding preferences can be irreversibly modulated by learning, in this case, by a form of Pavlovian conditioning.

Pavlovian conditioning can affect eating in other ways as well. Through little-understood processes, animals can learn about the different nutritional values of foods that taste and look different, for example. Animals seem to show a few specific hungers for nutrients necessary to relieve specific deficits. Hunger for salt is the best documented, but there may be specific hungers for some vitamins and amino acids, as well as learned preferences for tastes associated with foods that relieve some specific deficiency. And certainly the palatability of certain foods – the sweetness of sugar, the tastiness of a steak, and so on – tends to reflect their nutritional value. The correlation is not perfect, of course. Artificial sweeteners, absent during human evolution and devoid of metabolic benefit, taste like sugar, for example. The pharmaceutical industry is ever in search of good-tasting, but non-nutritious, substances to satisfy the popular wish to be both slender and gluttonous.

Some of the connections between taste and preference are built-in: sugar is attractive to every child and its attractiveness does not depend on experience. Specific hunger for salt seems to be largely innate (but see Kriebelhaus & Wolf, 1968). But other tastes – for spicy food, for example – are probably acquired (Rozin, 1976). All acquired preferences are direct effects of learning on motivated behavior. There are also modulatory effects that involve learning only indirectly. Most animals show circadian rhythms of eating. Rats prefer to eat during the night for example – although they will change if food is available only during the day (Jensen, Collier, & Medvin, 1983) – so this apparently built-in effect can be modified by learning. Even something as basic as the duration of a meal may be affected by conditioning processes: rats seem to learn when to quit eating a particular food, based on its nutritional value (see, for example, Booth, 1972; Davis & Perez, 1993).

These are all effects of learning on motivational processes. In these cases, learning is a complicating factor. For some reason there has been much less interest in the reverse effect, the
involvement of motivational processes in behavior that is supposedly “learned.” This is a pity, because motivational processes are simpler than learning processes, so that that accounting for learned behavior in motivational terms would simpler than explaining it in terms of learning.

Motivational factors like hunger are usually assumed to be constant or changing only slowly during “open-economy” learning experiments, so that their effects can be ignored. But this convention assumes that we know much more about “hunger,” and the dynamics of motivational processes in general, than we actually do. When we get down to the moment-by-moment details of feeding we find that motivational dynamics cannot be ignored, even in supposedly “pure-learning” experiments.

The Scope of “Motivation”

The term motivation covers both the processes that govern the allocation of time to different activities – feeding, drinking, sleeping, sex and so on – and the processes that allow organisms to vary the amount of time and effort they devote to a particular motivated activity. I discussed allocation from an optimality point of view in Chapters 2 and 3. If we look at feeding (say) in isolation, motivation refers to regulation of the amount of feeding on a given food, and also preference for one food type over others.

It is not always obvious which of these various hypothetical processes is operating in a particular situation. For example, suppose a rat is both hungry and thirsty. It eats for a while and then drinks. Does it stop eating because of regulatory feedback from the “feeding system,” or because of inhibitory signals from the competing “thirst system”? In plain English, does it stop eating because it is no longer hungry, or because it is now more thirsty than hungry? How do drinking and eating depend on one another? Fortunately, data and theory suggest that we can often assume independence: at any time, one motivational condition seems to dominate, so that if an animal is more hungry than thirsty, drinking occurs only when the hunger system permits (McFarland, 1974). There are obvious exceptions to this independence. Rats show different eating patterns in a novel environment than a familiar one, for example, presumably because of competition from vigilance activities (Wiepkema, 1971). Nevertheless, my working assumption is that under suitably controlled conditions, feeding regulation can be understood as a separate process. I will show how feeding can be understood in isolation, leaving until later the question of how feeding and drinking interact. Because different species can arrange these things rather differently, I will concentrate on the best-studied species (after man): the laboratory rat. Fortunately, the rat seems to be quite similar in many ways both to man and to many other species. And because I am ultimately interested in the effects of motivational processes on behavior in reinforcement-learning experiments with a fixed food type, I will focus on feeding dynamics, rather than feeding preferences.

This chapter is a summary of experimental results on the regulatory dynamics of feeding in rats. The next two chapters describe a model for the process.

Box 7.1: History of Motivation Research

Significant figures in unraveling the physiological and behavioral basis of motivation are the French physiologist Claude Bernard (1813-1878), the American physiologist Walter Cannon (1871-1945) and the American psychologist Curt Richter (1894-1988). Richter has summarized the history:

In 1859 Claude Bernard first described what he called the internal environment of the body, consisting largely of the body fluids, and showed that in mammals the properties of this internal environment ordinarily vary within fixed limits, variation outside of those ranges endangering life. He described many of the physiological mechanisms by means of which the body keeps these properties at fixed levels, and
pointed out that it is by virtue of the existence of these mechanisms that mammals are able to live and thrive under widely varying external conditions.

Cannon, in a long series of remarkable experiments, collected in 1932 in his book “The Wisdom of the Body,” not only confirmed Bernard’s concept but greatly extended it. Largely through his efforts this concept has become almost an axiom of modern medicine. Cannon speaks of a constant state or homeostasis. Thus he states: “The constant conditions which are maintained in the body might be termed equilibria. That word, however, has come to have a fairly exact meaning as applied to relatively simple physico-chemical states, in closed systems, where known forces are balanced. The coordinated physiological processes which maintain most of the steady states in the organism are so complex and so peculiar to living beings – involving, as they may, the brain and nerves, the heart, lungs, kidney, and spleen, all working cooperatively – that I have suggested a special designation for these states, homeostasis.”

Both Bernard and Cannon concerned themselves almost entirely with the physiological and chemical regulators of the internal environment. They showed, for instance, that when an animal is placed in a cold external environment and is consequently threatened with a decrease in body temperature, loss of heat is minimized by decreased activity of the sweat glands and constriction of the peripheral blood vessels, and more heat is produced by increased burning of stored fat and by shivering. These are all physiological or chemical regulators.

The results of our own experiments have shown that behavior... also contribute[s] to the maintenance of a constant internal environment (Richter, 1942/1976, pp. 194-5).

These behavioral regulatory mechanisms are the topic of this and the two following chapters.

BEHAVIORAL REGULATION OF FEEDING: EXPERIMENTAL RESULTS

Feeding is the prototypical motivated activity and the most obvious aspect of feeding is that it is regulatory. In this section, I describe experimental results that demonstrate regulation at different time scales.

Figure 7.1 illustrates body-weight regulation, a long-time-scale process. The figure shows the average body weight of a group of rats over time. Weight increases smoothly as the animals grow. At the point labeled “1” the group was split into two. Half were force fed (“a”): their weight increased above the normal growth line (“b”). The other half were placed on a restricted diet (“c”): their weight decreased below the line. At point “2” the animals were permitted to feed normally once again. The “a” rats eat less than normal, the “c” rats eat more, until both groups return to the curve of normal growth. The effects of almost anything that perturbs body weight – forced exercise, which usually reduces body weight, or diet-induced increased feeding (hyperphagia), which increases it, are two other examples – are reversed when the treatment is terminated.

The regulatory property of feeding has been described in several ways, with more or less theoretical overtones. It is said, for example, that body weight is “defended,” or “regulated” according to a “set point.” All I intend by “regulated” is that there are properties of the feeding system, like body weight, that return more or less to their initial value after a perturbation. The most neutral way to describe the regulatory property is in the language of Chapter 4, as
a stable fixed-point equilibrium. We don't need to commit ourselves at this stage to the existence of a physiological set point or to any particular form of negative feedback.

**Effects of Interruption**

Weight regulation is regulation over the very long term – days or weeks. Feeding is also regulated over much shorter periods, minutes and hours. In this case, of course, the thing regulated – the *regulated variable* – is not body weight, but some property of ingestion: eating rate (grams/hour of rat chow consumed) or energy-intake rate (calories/hour of food ingested). The simplest way to show regulation over a short time scale is to interrupt the normal pattern of feeding for a while and watch what happens when the opportunity to eat is restored. The result is shown schematically in Figure 7.2. The figure shows eating episodes (e.g., pellets ingested, laps of sugar-water) as vertical blips (fused together in the figure) across the top and cumulative food intake underneath. Feeding is interrupted during the period marked by the heavy horizontal line. The figure shows three things that are characteristic of feeding in rats, chicks, humans and many other animals:

(a) **Meals.** When ad lib. food is available, eating occurs in bouts, called meals. There has been much discussion in the feeding literature about just how to define a meal. As we'll see in a moment, this may not be the most useful question to ask. Nevertheless, the experimental data are reasonably clear: if you look at an event record of eating episodes (like the blips at the top of Figure 7.2), you can clearly see groups of blips (“meals”) separated by longer periods. If you plot a distribution of the times between eating episodes, it will show a peak at the typical time between eating episodes within a meal and a smeared tail at longer values corresponding to one or two modal between-meal times. A typical inter-lick-time distribution (plotted as a “survivorship curve,” which shows at each time the number of meals as long or longer than that time) is shown in Figure 7.3.

The conventional “meal” vs. “intermeal” designation is usually taken to imply just two underlying processes, but later work (Davis, 1989; Berdoy, 1993) suggests that at least three processes are involved (lower graph in Figure 7.3), which agrees better with typical descriptions of rat eating behavior: e.g., “Meal periods were typically composed of one or two bursts (Lucas and Timberlake, 1988, p. 261)”. The time between bursts within a meal presumably reflects the third process (middle line in the bottom panel of Figure 7.3). Obviously, if three or more processes are involved in the dynamics of eating, the “meal” loses its unitary quality and becomes less useful as an organizing concept.

---

86 The fixed point, the stable body weight, is changing slowly, of course, as the rat matures. In most species the stable adult body weight changes very slowly or not at all, however. In either case, the equilibrium properties at short-to-medium time scales are similar.
(b) **First-meal Effect.** If eating is interrupted for a period of hours, the effect is almost entirely on the first meal after eating resumes (the arrow in Figure 7.2). The noted French (*naturellement!*) student of feeding, Jacques Le Magnen, writes: “A food withdrawal from the beginning of the night for 2, 4, 6, . . . h, leads to a linear increase in the size of the first meal at the restoration of food access. Thus, the size of meal, unaffected by the duration of the pre-meal period of no eating in the ad libitum condition, becomes highly dependent on this pre-prandial interval of no eating when it is prolonged by the removal of food access. (1985, p. 22)” In Figure 7.2, meals last 11 time steps, but after the interruption there is a single extra-long meal of 19 TS; subsequent meals are again just 11 TS. I’ll call this the first-meal effect of interruption.

Le Magnen also points out an interesting contrast between normal eating and eating following a perturbation. Under free conditions, even though both meal length and intermeal interval show some variability, there is no correlation between intermeal interval and the duration of the following meal; that is, the rat is no more likely to eat a long meal after a long intermeal interval than after a short interval. (There is often a small correlation between meal length and the length of the following intermeal interval, however: Le Magnen & Tallon, 1966; Panksepp, 1973, 1976; see also Clifton, 1987). Rats are bit more likely to wait a long time before eating again after they have just finished an extra-long meal.) Nevertheless, when the intermeal interval is artificially prolonged (i.e., food is removed), a strong correlation emerges between the length of interruption and the length of the first meal after feeding is resumed, as illustrated in Figure 7.2. Thus, absence of a correlation between two variables (like meal length and intermeal interval) under free conditions, i.e., in the absence of any constraint, does not rule out the possibility of a causal relation between them.

(c) **Regulation.** The fact that there is a direct relation between the duration of the interruption and the size of the first post-interruption meal means that rats will partially compensate for an interruption of normal eating. In practice, animals compensate almost completely by this

---

87 As is well known (but often forgotten) in other contexts, just as absence of a correlation does not mean absence of causation, presence of a correlation under free conditions doesn’t (necessarily) mean presence of causation. The correlations between hospitals and sickness, between skirt length and changes in the stock market, between many aspects of diet and longevity are common examples of the dangers of inferring causation from correlation. Some correlations are causal some are not, but which is which? Direct experimental intervention is the simplest way to decide the issue.
first meal for the loss of food caused by the interruption. Some of the residual shortfall is made up slowly over many later meals, by shortening of the intermeal interval and some increase in meal duration.

These conclusions are supported by the results of other, more elaborate, experiments. For example, Lucas and Timberlake (1988)\textsuperscript{88} looked at free feeding in rats with different imposed minimum delays between each food pellet. Under free conditions, each rat could eat from a trough containing a single pellet. Removal of the pellet immediately triggered (via a photocell arrangement) the delivery of another pellet, in a potentially endless supply. Under the delay conditions, however, a delay of 16, 32, 64, 128, 256 or 512 s was imposed between removal of one pellet and delivery of the next. The experimenters looked at mean meal size\textsuperscript{89} (in grams), meal duration, and intermeal interval, as well as total intake, at each delay value. They found that food intake overall declined slowly with increasing interpellet delay. Total intake declined significantly only at the longest delays. They also found that the shortest delay reduced meal size by comparison with free eating, but longer delays caused no further reduction. Because meal size ceased to get any smaller as delay was increased from 16 - 512 s they concluded that “meal size is determined . . . by a positive feedback component with a relatively short time decay time probably based on gustatory stimulation from feeding . . .” Because intermeal interval decreased to (partially) compensate for reduced meal size, they inferred the existence of “a negative feedback component capable of integrating total intake across delays of up to 1 hour (Lucas & Timberlake, p. 259).”

**Short-term Regulation**

I have described feeding regulation across days (weight regulation); across hours (the effects of long interruptions); and across tens of minutes (the effect of interpellet delays). There is also regulation over seconds and minutes, i.e., within a meal. Most of the work on short-term regulation has been done with liquid diets, such as sucrose solution, where ingestion rate can be better monitored and controlled than with solid food. There are several ways to do these experiments. Ingestion can either be free or controlled. In the first case, the rat laps at a spout connected to a graduated cylinder of liquid food and thus controls both the rate (how fast he laps) and timing (when he laps) of food delivery. In the second case, ingestion is via an intra-oral infusion pump that fixes the rate, but not the timing, of food delivery. The interruptions can be in a single block, or repeated. The best dependent measure is simply the experienced time pattern of eating episodes. But (such are the conventions of the field) a summary measure such as meal size or length is often used instead of a comprehensive moment-by-moment record of ingestion. Here are two fairly typical studies that show how “within-meal” eating is regulated in the face of interruptions.

Seeley, Kaplan, and Grill (1993) describe two highly controlled experiments in which rats were infused intra-orally with 12.5% glucose solution at a rate of 0.75 ml/min. The infusion system detected rejection by the rat and stopped the pump immediately. Each rat was tested once a day. No eating was permitted in the 5 hours before each test. During a test, after the first rejection, infusion was stopped for 30 s. If the rat rejected the glucose solution again within 60 s after infusion resumed, the intake test was terminated. If not, this procedure was repeated until

\textsuperscript{88} These authors separated data from the day and night portions of the diurnal light-dark cycle, but this is irrelevant to the present argument.

\textsuperscript{89} Mindful of the uncertainties discussed in connection with of Figure 7.3, Lucas and Timberlake set a generous meal criterion: consumption of at least four pellets, followed by a pause of > 10 min without eating after a pellet became available.
the rat rejected twice within a 90-s period. The length and size of this single daily “meal” were defined over the time when infusion began and finally ceased.

In the two experiments, one, or several, interruptions were scheduled early during a meal, i.e., before the first rejection. Seeley et al. were interested in the effects on total session intake of interruption duration, of time of onset (whether the interruption occurred early or late within the meal), and of interruption pattern (whether the interruptions occurred in a block or were spaced). What were the effects of these different perturbations? In a word, none: the amount ingested was constant under all conditions. The rats in these two experiments defended their total per-meal intake perfectly.

In a second set of experiments, by Mook and Dreifuss (1986), rats maintained at 80% of their free-feeding weights were allowed to lap freely at a spout connected to a cylinder of saccharin solution. Saccharin is nonnutritive and moderately attractive to rats. They treat it like food, not water: “[T]he amount ingested increases with severity of food deprivation. . . whereas water intake declines (Mook & Dreifuss, p. 365-6).” Under free conditions, the rate of lapping by rats at the saccharin spout declined smoothly across a 30-min test session (open triangles in Figure 7.4). The striking result is that reducing lap opportunities by 50%, by interrupting lapping in alternating 15-s or 30-s periods (open and closed circles in the figure), made almost no difference to the total number of laps. The rats simply lapped faster in the reduced time available so as to maintain the overall rate of lapping approximately constant. In this experiment, lap rate, ingestion rate and meal duration were all simultaneously regulated. These experiments with saccharin show that there are short-term regulatory feeding processes that depend only on sensory feedback from the mouth. Since saccharin has no nutritive consequences, no negative feedback from the stomach or later in the digestive chain is involved.

**Feeding Physiology: Sensory Feedback**

*Sham feeding* is a popular surgical intervention that aims to separate the taste (orosensory) effects of food from effects later in the ingestional sequence, such as stomach distention, absorption from the gut and nutritional and metabolic effects. A fistula that diverts ingested food out of the body is implanted in the animal’s gullet, so that food never reaches the stomach. Ingestion under these conditions is assumed to be determined solely by taste factors. Mook and Dreifuss point out that “As a model of the orosensory control of ingestion, saccharin intake by hungry rats...parallels sham feeding in many respects. Like sham feeding, it is responsive to deprivation...Conversely, it is depressed by counter-experiments that impose postingestive controls,
just as sham feeding is...Like sham feeding, it fails to produce the normal behavioral correlates of satiety...Indeed, from a nutritional point of view a bout of saccharin ingestion is a case of sham feeding, one that requires no surgical intervention...Finally, in the case of saccharin (though not other sweet solutions), rats that are literally sham drinking, through open gastric cannulas, drink no more than when they are drinking normally with cannulas closed. (1986, pp. 365-6)

Thus, we might expect sham-feeding studies to show taste-determined regulatory properties just like the saccharin studies, and they do. For example, Davis and Smith (1990, cited in Davis, in press) studied rats, equipped with a sham-feeding fistula, lapping 0.8M sucrose solution. They looked at the rate of lapping across a 30-min period with the fistula open (sham feeding) and fistula closed (normal feeding). In both cases, lapping rate declines through the session, but the rate of decline is greater for normal feeding. But the fact that the first-sham curve does decline confirms our inference from the saccharin studies: there are inhibitory (i.e., negative feedback) effects on feeding from sensory factors alone. Comparison between normal and first-sham curves shows that additional negative feedback also comes from processes later in the chain from ingestion through absorption. Experiments in which rats learn to press a lever solely for intravenous food (e.g., Nicolaïdis & Rowland, 1977) also show negative feedback from later in the chain.

Davis and Smith gave a total of five sham-feeding tests in this experiment. They present a third set of data points showing the lapping-rate profile on the fifth sham feeding test: the fifth-sham curve declines more slowly that the first-sham curve described above. The effect of experience on sham feeding is evidently to reduce the decline of lapping through the session. Far from learning that sham-feeding is nonnutritive, the rats seem to want more of it. On the basis of this and subsequent experiments, Davis concludes that this difference between the effects of the first and fifth sham tests reflects Pavlovian conditioning: the first-sham animals inhibit drinking partly because of experience with normal feeding which produces anticipatory satiation. This experience dissipates across successive sham-feeding trials, so that feeding declines more slowly on later trials. These effects of experience are a complication we will have something to say about later.

Sham feeding allows researchers to isolate orosensory factors from effects further down the chain. Other techniques, such as the pyloric cuff, nonnutritive or bad-tasting dietary admixtures, drug treatments and brain manipulations (stimulation, lesions), begin to answer additional questions about physiology: What are the effects on eating of taste and gastric factors? Of food-absorption rate? Of stores in liver? Of the levels of circulating metabolites? These are important concerns for a full understanding of feeding. But my concern in this chapter is much more limited: just with feeding dynamics. There are two implications of the sham-feeding studies for dynamics: (a) feeding regulation operates at several time scales, short (orosensory factors) and longer (gastric and other factors), and (b) some associative learning may be involved in the inhibitory feedback from sensory factors.

**Incentive, Palatability and Deprivation**

Feeding is affected by the taste of food, its nutritive value and the animal's state of food deprivation. These factors are sometimes labeled with terms like “palatability,” “incentive” and “deprivation.” “Incentive” and “palatability,” like “hunger” and “satiety,” are terms derived from shared properties of our own subjective experience. They do not necessarily correspond to identifiable or unitary properties of the feeding mechanism. Nevertheless, they do serve to label important aspects of feeding.

Incentive refers to the “motivating value” of a particular food type. Incentive is under many conditions correlated with nutritive value. To avoid confusion we’ll try to stick to objec-
tive terms like “concentration” or “caloric content.” But even a simple incentive variable like nutrient concentration affects the organism in two ways: directly (via taste differences, and differences in nutritional uptake) and, after some experience, through learning: the organism learns to associate a particular taste or food source with certain nutritional consequences. Sometimes it is difficult to separate these two kinds of incentive: the rat laps faster the more concentrated the sucrose solution. Is this because of an innate linkage between the sweeter taste and greater caloric content, or does the rat have to learn the connection? Probably both are involved: there is certainly a “default” preference for sweetness, but it can undoubtedly be modulated somewhat by experience.

Rats do lap faster at sweeter sugar solutions, at least at first. Figure 7.5 shows number of laps during a brief (3-min) exposure period to five or six concentrations of four different sugars. The function relating lap rate to concentration is positive and almost linear over much of the range for these sugars.

Faster lapping of a more nutritious solution implies a failure of regulation: the animal should ingest less of the more nutritious solution if it is to maintain a constant caloric intake. Incentive effects seem to be anti-regulatory. Regulation reappears at longer exposures, however, because a low (initial) rate of lapping (of a mildly palatable solution) usually persists longer than high-rate lapping at a palatable solution. Figure 7.6 shows an example: “The initial rate of the more concentrated [0.8M] solution is greater than that of the less concentrated [0.05M] solution, but the rate of decline is steeper (Davis, 1989, p. 108).” Consequently after a long enough period the rat will ingest almost equal amounts of sucrose from both solutions. I describe a way to combine regulatory and incentive factors in the next chapter.

Eating solid food under free conditions also shows regulation. Johnson, Ackroff, Peters & Collier (1986), compared eating patterns with food of different caloric densities. They note that “rats maintained a constant daily caloric intake by eating more frequent, larger meals of the lower density foods. . . Higher-calorie foods were consumed at a faster rate within meals than were lower-calorie foods.” Note that most of the regulation occurs via a change in meal duration, rather than meal frequency. Johnson et al. conclude that “caloric intake may be regulated over a time frame of several meals rather than on a meal-to-meal basis. (Johnson, Ackroff, Peters & Collier, 1986, p. 929; see also Hirsch, 1973, for similar data in guinea pigs)” The issue here is the time window over which regulation occurs. The liquid-diet data in
Figure 7.6 show that if the rat is undeprived and the intake period is short enough, caloric intake is not regulated. Johnson et al.’s data in Figure 7.7 show that if the animal is allowed to control meal duration and frequency, intake is regulated.

**Palatability** refers to the properties of taste, apart from nutritional consequences. The results of sham-feeding experiments, experiments with saccharin or other artificial sweeteners, and experiments with bad-tasting adulterants like quinine, display the effects of palatability by itself, separate from incentive. Since taste is sensed by the mouth (in humans and rats), taste factors have their greatest effects early in the ingestional sequence.

**Deprivation** can be objectively defined, in terms of body weight relative to a free-feeding norm or by hours since the last meal. Under free-feeding conditions, increasing food deprivation causes rats (eating solid food) to eat in longer bouts separated by shorter interbout intervals. The overall number of bouts per unit time remains more or less constant (Wiepkema, 1971) but overall food intake rate increases with deprivation. Notice that there is some contradiction between the effects of long-term and short-term deprivation. Short-term deprivation (hours) has an effect mainly on the first meal after resumption (Le Magnen, 1985; Levitsky, Faust & Glassman, 1976); long-term deprivation (days) may have an effect on the length of several meals and on intermeal interval.

How can all these effects be explained? In particular, can they be explained by purely regulatory mechanisms, or is something more required. The next two chapters attempt to answer this question.

---

**Figure 7.7** Caloric regulation under free conditions. Mean daily intake in calories (*top*) and grams (*bottom*) of four diets differing in caloric density by free-feeding (N=8) and foraging (N=8) rats. (From Johnson, Ackroff, Peters & Collier, 1986, Figure 2.)
Chapter 8

MOTIVATION, II: A MODEL FOR FEEDING DYNAMICS

The ultimate objective of model-making in behavioral neuroscience is to understand the relations between neurophysiological structure and function, on the one hand, and the behavior of the intact organism on the other. The obvious route to this goal is to construct models out of ingredients drawn from the known physiological facts. We know, for example, that taste of food alone has some inhibitory effect on further eating; some behavioral data suggest that taste may also have a transient positive effect (l’appetit vient en mangeant: the "salted-peanut effect"). That stomach fill has an inhibitory effect on further eating is well established. We also know that the body's metabolic reserves, in the liver and elsewhere, exert modulating effects. Literal models – models whose elements are directly identified with these known physiological factors – therefore have unique face validity.

Much can be learned from literal models. If the models fail, we will know that something is wrong or missing: failure provides clues for future physiological research. On the other hand, if they succeed, we have proof of the essential completeness of physiological understanding. Despite their surface plausibility, and despite the fact that this has been the dominant approach to understanding feeding mechanisms, literal models have not so far been very successful in describing the dynamics of feeding behavior.

One reason is that the feeding system is probably very complicated, from a neurophysiological point of view. I discussed the issue of complexity in Chapter 1. But the main problem with using physiologically incomplete models to help us understand a complex system is that if even one element is omitted or wrong, the model may fail massively. The elements of a complex system cannot just be added together so that if we have 90% of the elements our predictions will be 90% correct. They may be 100% wrong. DDT kills medflies, but a campaign of DDT spraying may actually increase the number of medflies if it kills medfly predators more effectively than medflies themselves. What incomplete models show is often interesting, but what is left out may be vital. Nothing critical can be omitted from dynamic models, if we hope to understand the effects of behavioral experiments. But in fact, our understanding of the physiological underpinnings of feeding behavior is incomplete. The list of taste factors, stomach fill, liver stores and circulating metabolites, etc., does not exhaust the physiological factors involved in feeding. As Cannon pointed out: “The coordinated physiological processes which maintain most of the steady states in the organism are so complex and so peculiar to living beings – involving...the brain and nerves, the heart, lungs, kidney, and spleen, all working cooperatively...” A model, no matter how “biologically plausible,” that limits itself just to known physiology is almost certainly incomplete.

If this view is correct, then at this stage of the game the best guide to modeling feeding behavior may be not physiology, but parsimony and the behavior itself. The proper question may be not “How does known physiology explain the properties of feeding behavior?” but “What is the simplest dynamic model, consistent with but not limited by known physiology, that can explain the properties of feeding behavior?” It would be foolish for our models to ignore what is known about feeding physiology. It would be equally foolish to be limited by what is known when what is known is known to be limited. Eventually, literal and behavioral models may converge on the actual physiological processes that underlie behavior. But we cannot assume convergence until the model deals adequately with the full range of behavioral data – a challenge that so far few literal models have accepted much less met.
BEHAVIORAL REGULATION OF FEEDING: THEORY

Why do animals eat? The functional answer (in Tinbergen's, 1974, sense of “function”) is obvious: they eat to maintain themselves in nutritional balance. It isn't quite as obvious why they eat in meals, why they eat more at particular times of day than others, or why they often eat more than they seem to need, but the functional questions are nevertheless relatively straightforward. The explanation for why animals eat in meals seems to be: (a) they eat much faster than they metabolize \(^{90}\), hence they don't need to eat all the time. (b) So long as food is available, it's more efficient to eat in bursts than slowly throughout the day; but (c) the potential size of the burst is limited by internal holding capacity (i.e., stomach size, for short-term storage, fat reserves for long term). The holding capacity, in turn, is limited by its cost, in terms of metabolic maintenance and reduced mobility.

Rats eat mostly during the dark period of the diurnal cycle, presumably because they are then less subject to predation, but circadian rhythms of eating are not inflexible. Rats will eat during the day, rather than at night, if there is some benefit to doing so (Panksepp & Krost, 1975). But the existence of a “default” pattern obviously makes sense, given the risks under natural conditions. And animals may eat to excess for the same reason that a driver tops up a half-full gas tank: because food is “cheap” (i.e., a highly nutritious food source becomes available) or because a long dry stretch is in prospect. Rats repeatedly cycled through periods of starvation alternated with periods of ad lib. food gradually increase their intake during ad lib. periods, for example (Brownell, Greenwood, Stellar, & Schrager, 1986). They also come to metabolize more slowly.

These adaptive considerations are hard to quantify, of course, and there is always argument about the extent to which any phenotypic property, behavioral or morphological, is really optimal \(^{91}\). But the adaptive function of natural eating patterns seems to pose no special problem.

Mechanisms

The dynamic mechanisms that drive feeding are not so simple to discover. In the 1970s, there was great interest in quantitative models for feeding behavior based upon what was then known of the underlying physiology – i.e., in literal models. Several such models were proposed in collections edited by McFarland (1974), Booth (1978), and Toates and Halliday (1980; see also Toates & Evans, 1987; Guillot & Meyer, 1987). In recent years, interest in models of this type has waned somewhat, perhaps for the reasons just discussed. In any event, attempts to make literal models for the dynamics of feeding were more or less abandoned in the 1980s. Nevertheless, experimental data on the dynamics of feeding have continued to accumulate, as I de-

\(^{90}\) This is true for rats and people; it is not so true for many herbivores or for small animals like shrews and hummingbirds with very high metabolic rates.

\(^{91}\) See the comments on functional explanation in Chapters 2 and 3, as well as the discussions of adaptationism by Oster & Wilson (1978), Maynard Smith (1978), Staddon (1983) and Orzack & Sober (1994; see also Brandon (1990) and Brandon & Rausher, 1996).
scribed in the previous chapter. The time is ripe to have another go at the problem, without the “literal” baggage.

**The Davis, Collins & Levine Model**

My strategy is to begin with the simplest literal model and show how behavioral data require modifications, adding at each stage as little as possible. A particularly simple model was proposed by John Davis and his associates (Davis, Collins & Levine, 1976; Davis & Levine, 1977). It is a good starting point, because it is built of elements used in all other models. It was designed to explain specific behavioral data — the pattern of licks by a hungry rat given free access to sucrose solutions of different concentrations — in terms of a specific negative feedback signal, distention of the gut. “This signal we believe arises from the activation of tension receptors in the wall of the small intestine and is activated whenever the rate of ingestion and the rate of delivery of chyme from the stomach to the duodenum exceeds the rate of absorption from the intestine causing a net accumulation of fluid in the intestine.” (Davis, Collins & Levine, 1976, pp. 395-6). The physiological assumptions are less important for our purposes than the formal properties of this model, which is an inflow-outflow system of the familiar leaky-bucket type.

**Figure 8.1**, Panel A, shows the original version of this model. The basic idea is exceedingly simple: (liquid) food comes in, at a rate determined by palatability. This is indicated by the direct feedforward path from “gustatory input” to “drinking-rate output” in the figure. At the same time, the ingesta flow into the gut at a rate determined by the rate of eating and rate of absorption from the gut. The model assumes a constant rate of absorption so that the net rate of fill, \( x(t) \), is proportional to ingestion rate. The “gut contents” integrator in Figure 8.1A is thus a simple accumulator. The net accumulation in the gut \( y(t) \) (in Figure 8.1A) exerts an inhibitory effect on ingestion: it’s simply subtracted from the gustatory input, so that the net ingestion rate is proportional to the difference between gustatory input and accumulation. When the inhibitory effect of accumulation equals gustatory input, the net “drinking-rate output” is zero. Formally, in the discrete-time notation used in previous chapters:

\[
x(t) = G(t) - y(t),
\]

and

\[
y(t+1) = y(t) + \beta x(t),
\]

where \( x(t) \) is eating (drinking) rate at time \( t \), \( G(t) \) is gustatory input (usually a constant) and \( y(t) \) is gut contents; \( \beta \) is a parameter that represents the fraction of inflow, \( x(t) \), in each time step that actually accumulates in the gut.

Davis et al. tested this model by comparing the eating-rate function with the eating (drinking) rate vs. time function following exposure to a given sucrose solution. The prediction is for an exponential decline in eating rate (Figure 8.1B), which matches the data well. They also looked at the effects of a manipulation designed to affect rate of absorption from the gut (hence rate of accumulation in the gut) on the eating function. Here the match of model to data is not so good.

In this simple form, the model has an obvious lack: What happens when \( y(t) \) reaches the equilibrium value and feeding ceases? Only rate of accumulation in the gut (represented by the \( \beta \) parameter) is specified in the model. Without some provision for absorption from the gut, the model cannot say anything about when eating should resume after repletion. Absorption rate can be incorporated in several ways. Perhaps the simplest is to assume that amount absorbed is proportional to gut contents. This just means adding a single parameter to **Equation 8.2**:

\[
y(t+1) = \alpha y(t) + \beta x(t),
\]

where \( 0 < \alpha < 1 \).
which says that in the absence of ingestion \( (x(t) = 0) \), gut contents, \( y(t) \), decline by a constant fraction \( 1-\alpha \) during each time step. This change converts the simple accumulator into the familiar two-parameter leaky integrator discussed in earlier chapters. The change still permits the model to show an approximately exponential post-deprivation repletion function, so long as \( \alpha \) is close to unity. But now eating never ceases: it simply declines to a low, steady level equal to the net absorption rate. Real animals, of course, cease eating entirely at the end of a meal and then resume after an intermeal period. Thus, a major flaw in this and several other simple feeding models is that they fail to account for the most striking feature of eating under free conditions: eating in meals. How can eating in meals be explained?

One possibility is suggested by the data of Lucas and Timberlake, discussed earlier, on the effects of enforcing delays after the eating of each pellet. They found that as soon as a short (16-s) delay was introduced, the animals ate smaller meals — but additional delays caused no further reduction in meal size. They concluded that there must be a taste-induced positive-feedback process that extends meal duration but dissipates if a delay is enforced between pellet deliveries. Both this process, and the negative feedback process depicted in Figure 8.1 and Equations 8.1-8.3, are already familiar from earlier chapters. The negative-feedback process corresponds to a feedback version of the feedforward habituation model we have discussed at length. The positive feedback is a version of sensitization.

Figure 8.2 shows one way that positive feedback can be added to the simple Davis et al. model. “Integrator 1,” with an inhibitory (gray arrow) effect on gustatory input, is the same as the integrator in the previous figure. But two new elements have been added: a second, excitatory, integrator (“Integrator 2”), whose output is added to the gustatory input (solid arrow), and a threshold (circle in the center), usually set equal to zero so that neither integrator gets a negative input.

The prethreshold sum of gustatory input minus inhibitory integrator plus excitatory integrator is just what we have been calling reflex strength; ingestion rate is the above-threshold value of reflex strength.

The equations for this modified system are:

\[
x(t) = G(t) - y_I(t) + y_E(t),
\]

if \( x(t) > \theta \), 0 otherwise

\[
y_{I}(t+1) = \alpha_{I} y_{I}(t) + \beta_{I} x(t),
\]

\[
y_{E}(t+1) = \alpha_{E} y_{E}(t) + \beta_{E} x(t),
\]

where \( \theta \) is the threshold (usually 0), \( \alpha_{i} \) are the time constants and \( \beta_{i} \) the increment-rate parameters of the two integrators.

If \( \alpha_{I} > \alpha_{E} \) (i.e., the inhibitory integrator is slower than the excitatory, as the data suggest) and \( \beta_{I} \) and \( \beta_{E} \) are appropriately chosen, this system will indeed show eating in bouts (i.e., “meals”). An example of the steady-state behavior of this system is shown in Figure 8.3. The
above-zero part of the reflex strength record corresponds to eating, the below-zero part to the intermeal period.

But the behavior of the system still differs in many ways from real eating patterns. The major defect is that this system is not regulatory. If the gustatory input (i.e., taste) is increased, overall eating rate increases almost in proportion. This is not what happens (Wirtshafter & Davis, 1977). As I pointed out earlier (Figure 7.6), animals drink a more palatable diet faster, but their rate of drinking also declines faster. Overall, rats maintained on more palatable diets eat only slightly more than rats maintained on less palatable diets. The system in Figure 8.2 also fails to regulate if a delay is imposed in between each lap: neither meal size nor total intake is defended, even if the delay is small. The first meal after a long interruption is somewhat smaller than usual. Clearly, this is a poor model for feeding regulation as a whole.

A Multi-Compartment Model

A model based on the dynamics of a single meal seems to have difficulties accommodating regulation over longer time intervals. Figure 8.4 shows a different approach, using the same elements, that begins with long-term regulation (Staddon & Zanutto, 1997). The basic idea is straightforward. When food is ingested it passes through several stages of processing: orosensory, gastric and finally metabolic. These stages can be illustrated by compartments, with metaphorical “fluid” flowing from one to another, as shown in the figure. (This picture is intended only to illustrate the dynamics of the model; I don’t propose to identify every compartment with a specific physiological process, for the reasons discussed earlier.)

At each stage, the effects of ingested food on the feeding system decline with time after the stimulus (food or the processed consequences of food) has ceased. Probably the peripheral (orosensory) effects are the most transient and the metabolic (energy expenditure) are

---

92 Peter Killeen (1988) has explored a similar metaphor for B. F. Skinner’s prematurely abandoned concept of the reflex reserve.
the slowest. In short, the simplest assumption is that the compartments and flows through them are more or less as suggested in the picture: “food in” is a high-rate process (eating is much faster than metabolizing), so that the level of the first compartment falls most rapidly; “energy out” (metabolism) is slow, so that the level of the last compartment falls most slowly.

My colleagues and I have explored a number of ways to implement this multi-compartment approach. Making it regulatory is not a problem—the fluid-flow metaphor shows pretty clearly that this system tends towards a stable fixed-point equilibrium. Look again at Figure 8.4. Suppose we pour in fluid (“food in” = eat) only when fluid level in the last compartment falls below a fixed threshold. It should be intuitively obvious that so long as inflow rate is not too large (so the top compartment overflows) or too small (lower than “energy out”), the average rate of inflow—i.e., total amount of fluid per unit time, averaged over a number of periods of inflow and no-inflow (molar ingestion rate)—will equal rate of outflow. If the level in the last compartment is constant (because of the threshold), rate of outflow will be constant: hence rate of inflow will also be constant. In the steady state, molar ingestion rate will equal outflow rate and be more or less independent of the local inflow rate.

Most important, this model also solves the eating-in-meals problem, because the effect on the fluid level in the last compartment of “food in” to the first compartment is delayed. More on how this works in a moment.

Although it obviously regulates at the molar level, it is not obvious that a system like this will, like real feeding, regulate at different time scales: within a meal, following short interruptions and after lengthy deprivation. To evaluate all the possibilities, we need to be more formal.

Figure 8.5 shows a more formal version of the hydraulic model in Figure 8.4, the cascaded-integrator (CINT) model (see also Figure 5.7). The compartments in Figure 8.4 are replaced by leaky integrators linked in a cascade. I have discussed a slightly different type of integrator cascade already, in connection with habituation (in Chapter 6). The habituation cascade is different, because units are linked by via thresholds, making the system as a whole nonlinear. The feeding cascade involves only a single threshold. Otherwise the integrators are connected directly (this part of the system is linear): the output of integrator \( N \) is the input of integrator \( N+1 \), etc.. The time constants are faster on the periphery, slower centrally: \( \alpha_1 < \alpha_2 < \alpha_3 \) (although this may not be essential). For the eating model, there is a reference point or threshold, represented by a fixed value, \( \theta \). The strength of the tendency to eat, “eat command,” is simply the difference between the weighted sum of integrator states and \( \theta \).

Formally, the cascaded-integrator model works like this. The tendency to eat (reflex strength, in our earlier terminology) is simply the difference between a reference value, \( \theta \), and an aggregate satiation signal, \( V_s \), which is the weighted sum of the states of the integrators:

\[
V_s = w_1 V_1 + w_2 V_2 + \ldots + w_N V_N = \sum_{i=1}^{N} w_i V_i.
\]  (8.7)
Physiologically, this assumption reflects the well-established fact that the tendency to eat depends not only on the state of bodily deprivation (i.e., events late in the cascade), but also on taste factors and stomach fill (i.e., events at the beginning and in the middle of the cascade: see Mook, 1987). The larger the value of $V_S$, the more satiated the animal and the smaller the tendency to eat. The $w$ values later in the chain are larger than those earlier (Figure 8.4 is an extreme example: only the last integrator in the chain contributes to $V_S$). Why? To produce eating in meals, as I explain in a moment.

To translate $V_S$ into an eating command (response strength) we need a reference, $\theta$: when $V_S$ falls below $\theta$, eating begins and the tendency to eat is proportional to the difference between $V_S$ and $\theta$. Thus, reflex strength $= \theta - V_S$.

$\theta$ resembles the set point that figures prominently in most discussions of feeding regulation, but $\theta$ is different in three respects: (a) It corresponds to no single physiological variable, such as blood glucose level or the level of fat stores. $\theta$ is a purely theoretical construct, which may reflect a whole constellation of physiological variables, represented by the $V_i$ terms in Equation 8.8. (b) $\theta$ doesn’t necessarily correspond to a fixed settling point for body weight. As we will see, different feeding regimens can lead to different steady-state weights, even though $\theta$ is constant. (c) Moreover, the system as a whole shows regulation of both long-term and local eating measures. Conventional set-point models would have to assume separate set points for variables at different time scales, but the cascade model needs only one fixed value, $\theta$.

How does this system handle eating in meals? Given a single threshold and a reflex strength, $\theta - V_S$, that includes the states of all the integrators in the chain (i.e., $w_i = 1$ for all $i$ in Equation 8.7), we again encounter the problem that the system will simply eat at a steady rate as $V_S$ cycles rapidly below and above the threshold, $\theta$. There are at least three ways to deal with this problem:

1. Introduce a positive-feedback loop, as already described. Each eating episode then charges up a slow-acting cascade of inhibitory integrators (the satiation signal) and also energizes a rapid-rising and fast decaying excitatory integrator (sensitization). In our simulations we have found that such a scheme may play a small role, but is not necessary to account for most of the data on feeding dynamics.

2. There is a simpler, but possibly less biologically plausible, method: two thresholds, a “start” threshold and a second, “stop” threshold. The idea here is that eating begins when $V_S$ falls below a “start” threshold, $\theta_A$, as in the one-threshold model; but once eating has begun, $V_S$ must rise above a second, higher, “stop” threshold, $\theta_B$ before it ceases. This version works surprisingly well (cf. Booth, 1978; Guillot & Meyer, 1987; Staddon, 1988b), but there are at least three objections to it: (a) The definition of an eating bout is arbitrary. In our simulations, a transition from not eating, in time step $t$, to eating, in time step $t+1$, triggers a change in the effective threshold from $\theta_A$ to $\theta_B$ which is arbitrary because it depends on the size of the time step. If the time step is changed, a completely different pattern of behavior may result. Moreover, (b) the two-threshold approach takes no account of the number of time steps in between eating bouts: two eating episodes separated by
one time step cause the same threshold change as two episodes separated by two or one hundred, which is not plausible. (c) The behavior is too perfect: every meal is uninterrupted, for example. The two-threshold assumption cannot generate three-state inter-eating time distributions like those in Figure 7.3.

3. The third possibility is also the simplest: to use as our satiation signal not the sum of all the integrator states, \( \sum V_i \), where \( i = 1, N \), but the sum only of the states of later integrators in the cascade, e.g., \( w_1, w_2 = 0; w_3 = 1 \), for a 3-unit cascade. When the weights of early integrators are low relative to the weights of later integrators in the chain they serve mainly as delays. As I explain in a moment, the accumulated delay is sufficient to produce eating in meals and does not interfere with useful regulatory properties of the cascaded system.

**Equations.** For the simple case, we assume that eating is all-or-none: the system eats at a fixed rate during a time step if the satiation signal is below threshold, and not otherwise. The size of each eating input reflects the satiating value (SV) of each bit of food. “Satiating value” is presumably related to taste, texture, bulk and nutrient value; these properties probably have separable effects on feeding, but it is simpler at this stage not to distinguish among them. The equations for this system under free-feeding conditions are:

\[
X(t) = \phi(t) \text{ if } V_S < \theta; 0 \text{ otherwise, (8.8)}
\]

\[
V_i(t+1) = a_i V_i(t) + b_i X(t), \quad 0 < i < N, (8.9)
\]

\[
V_i(t+1) = a_i V_i(t) + b_i V_{i-1}(t), \quad 0 < i < N, (8.10)
\]

where \( X(t) \) is the SV during each time step (i.e., the food actually eaten), \( \phi(t) \) is the SV of each bit of food available (controlled by the experimenter), \( V_i \) is the state of integrator \( i \), \( a_i \) is the time constant of integrator \( i \) and \( b_i \) is the input weight. Equation 8.8 says that the system eats the food available (size: \( \phi(t) \) in each time step) only if the satiation signal, \( V_S \), is below threshold, \( \theta \). **Equation 8.9** says that the state of the first integrator, \( V_i(t+1) \), is determined by its previous state, \( V_i(t) \), and the SV during that time step, \( X(t) \); **Equation 8.10** says that the state of an integrator later in the cascade, \( V_i(t+1) \), is determined by its previous state, \( V_i(t) \), and the state of the preceding integrator in the chain, \( V_{i-1}(t) \).

**Impulse Response.** The simplest way to understand the behavior of this system is to look at its response to a single eating episode (“bite”). The time profile of the response to an isolated, brief input is termed the impulse response of the system. Since the equations for \( V_S \) are linear, all its properties follow from this function. The impulse response for a five-integrator system, with \( V_S \) equal to the sum of \( V \)-values for the last three integrators (i.e., \( w_3, w_4, w_5 = 1 \), others = 0), is shown in Figure 8.6. The function is unimodal, and the maximal effect of a single piece of food is delayed by an amount determined by the time constants of the first two integrators. A func-

---

**Figure 8.7** Top panel: Steady-state eating pattern for the system illustrated in Figure 6.12, with an eating threshold, \( \theta = 1 \), and a “bite” (SV) size, \( \phi(t) = 1 \). The wavy line is the satiation signal (i.e., \( \Sigma V_i \), \( i=M,N \), for \( M=3, N=5 \)). The blips at the bottom show eating episodes. Bottom panel: Steady-state eating pattern for a bite size, \( \phi(t) = 0.5 \).
tion with this form is consistent with what is known of the delayed effects of taste and nutritive effects of feeding (see, for example, Panksepp’s, 1976, literal feeding model).

Meals. It is easy to see intuitively that this system will produce eating in meals. The process works like this. Let’s suppose that \( V_S \) has just fallen below threshold, so that the system begins to eat. Because of the lag in the impulse response, this eating will not have an immediate effect on \( V_S \), which will continue to fall for a while. Eating may therefore continue for several time steps before \( V_S \) begins to increase and several more before it increases enough to exceed the threshold and suppress further eating. The lag in the impulse function will cause a further increase in \( V_S \) before it begins to fall again; when \( V_S \) finally falls below threshold, eating resumes.

The steady-state pattern of eating shown by this system is illustrated in Figure 8.7 (top panel), which shows 50 time steps after a stabilization period of 900 steps. As you can see, the average meal length is between three and four time steps, for a “bite” size (SV) of one. The bottom panel shows the effect of reducing the SV to 0.5: the meal size increases to six or seven time steps, which means, of course, that the meal size in terms of amount eaten is approximately regulated. Overall intake is also regulated, because meal frequency does not change. Meal duration is not regulated, but in this simulation the system has no control over the amount ingested in each “bite” so cannot simultaneously regulate both meal duration and meal amount. The situation resembles the controlled-infusion-rate-sucrose solution procedure of Seeley et al. (1993): the model can only turn ingestion on and off; it has no control over the rate of ingestion in the “on” state.

As we saw earlier, experiments show that when food is made less nutritious, meal length increases proportionately, but meal frequency doesn’t change (see Johnson et al., 1986). Thus, the doubling in meal size caused by halving the satiating value (“bite size”) of each eating episode shown by the model in Figure 8.7 matches real eating.

Figure 8.8 shows a phase space representation of the system (i.e., model plus food schedule). The system state is defined by five \( V \) values, one for each integrator, so we cannot diagram it. But we can plot the sum of the first two integrators (i.e., the “lag” integrators) against the sum of the last three (the satiation signal), which is done in Figure 8.8. As you can see, the system settles down to a periodic pattern, which is a projection of the periodic attractor that causes the meal patterns shown in the previous figure.
Effects of Interruption. Figure 8.9 shows the effects of deprivation and interruption. The system begins with all V values set equal to zero – maximal deprivation. When eating is allowed to begin, the satiation signal increases almost linearly and eating occurs in a long initial meal (A in the figure), after which eating in the regular meal pattern resumes. A brief period (20 time steps) when eating is interrupted (dashed lines) is followed by a single long meal and resumption of the regular pattern. The effect of a brief interruption is precisely the “first-meal” effect that Le Magnen (1985) has described. The effect of a substantial interruption is constant eating in a long meal, rather than the exponentially decreasing rate of eating that Davis et al. describe (see Figure 8.1B) – but this is because of the “all-or-none” eating assumption in this simple version of the CINT model.

The effects of interruption resemble the “binge eating” so characteristic of anorexic individuals and both may reflect the same property of regulation of this type (Box 8.1).

Box 8.1B: Binge Eating

Eating disorders are common in the U.S., especially among young women:

About 3 percent of young women have one of the three main eating disorders: anorexia nervosa, bulimia nervosa, or binge-eating disorder… Binge-eating disorder is a newly recognized condition featuring episodic uncontrolled consumption, without compensatory activities, such as vomiting or laxative abuse, to avert weight gain …Bulimia, in contrast, is marked by both binge eating and by compensatory activities. Anorexia nervosa is characterized by low...
body weight (< 85 percent of expected weight), intense fear of weight gain, and an inaccurate perception of body weight or shape.


There are substantial individual differences among sufferers from this complex of conditions, but two features common to most are restriction of food intake (anorexia) and occasional bouts of excessive eating (binging). People with anorexia wish to be thin, even if they already are, and make strenuous efforts to restrict their food intake: “an extreme need to control eating is the central feature of the disorder.” (Fairburn, Shafran & Cooper, 1999) How could this be accomplished within the cascade eating model? The most obvious locus of action is the eating threshold. If the threshold is lower, the satiation signal must decline to a lower value before eating begins.

Figure 8.1B shows the effect of lowering the satiation threshold on (a) total amount ingested per unit time and (b) the pattern of eating. The left panel shows the “normal” pattern: frequent eating episodes of medium length. The right panel shows the effect of reducing the effect of the satiation threshold: less frequent, smaller eating episodes and a reduction in average intake. This result is typical and depends little on the number of stages and parameter values in the model.

So far, the model does not predict the “binging” pattern that is so characteristic of anorexia. However, maintenance of the pattern shown in the right panel requires that the threshold be permanently maintained at its low, abnormal level. I have not discussed what determines the threshold, but it has some direct relation to metabolic needs: the lower the body weight, the higher we may expect the threshold to be. As body weight declines, the threshold should increase. Hence, maintaining it at an abnormally low level surely becomes more and more difficult for the anorexic individual. If there is a lapse, if, that is, the threshold is allowed to rise to something like its normal level, the result will indeed be a very long eating bout, a binge. The effect will be much the same as a forced interruption of feeding in the normal individual: an especially large “first meal”. Thus, binging may represent a natural “failure mode” of this kind of regulatory system.

\[ \theta - V_S \]

Regulation: Short and Long Term. If the initial long meal in Figure 8.9 is briefly interrupted, the meal resumes after the interruption and its total size remains the same – exactly duplicating the results of Seeley et al., described earlier: hungry rats “defend” the size of their first meal following a long (5 hours) period of deprivation. The CINT model will also duplicate the data of Mook and Dreifuss once we relax the all-or-none eating rule and allow eating rate (i.e., rate of ingestion of a sucrose or saccharin solution) to vary proportional to the difference between the threshold and the satiation signal (\( \theta - V_S \)).

In retrospect, none of these experimental results is very surprising. Seeley et al. deprived their animals for five hours before the experiment; even a 38-min interruption (their longest) during the initial meal.
adds relatively little, proportionately, to the state of deprivation with which the animals began the experiment. The Mook and Dreifuss animals were at 80% of their free-feeding weights, which implies an even longer deprivation period. In neither case was initial deprivation significantly increased by the local interruptions during the first meal. Given that initial-meal length will be determined largely by the initial state of deprivation of the animal, it is not surprising that the length of the initial meal was largely unaffected by the interruptions. Essentially any dynamic process capable of producing both molar regulation and eating in meals should be able to duplicate the results of these experiments.

The CINT model regulates pretty well over the long term. Figure 8.10 shows a cumulative eating plot before and after a medium-duration (50 time steps) interruption. The dashed line is an extrapolation of the steady-state eating rate. As you can see, most of the loss caused by the interruption is made up by the first meal following the interruption (the “first-meal effect” discussed earlier). If the interruption is short enough, the long first meal makes up the whole discrepancy, but beyond a certain period, the discrepancy is never made up. If rats are prevented from feeding for 24 hours or more, they also never make up the lost food (Levitsky et al. 1976; I take a more careful look at these data in the next chapter).

Figure 8.11 shows how overall eating rate is affected by the satiation value of each piece of food. When each bit of food is less than about 0.4, the system “eats” during every time step, so that eating rate is proportional to SV. Above that point, eating rate is regulated at about 34 per 100 time steps, with a slight tendency for the regulated level to increase with SV: given a highly nutritious diet, the model will slowly “gain weight”.

The CINT model also regulates in the face of imposed delays in between pellets, but not perfectly (cf. Lucas & Timberlake, 1988). The heavy line in Figure 8.12 shows the effect of increasing the time between permitted “bites” (interpellet interval, IPI) from 1 to 128. Over a range of delays from 0-128 s, overall eating rate varies only about 10 percent; thereafter the decline is more substantial. The open squares are data from Lucas and Timberlake with one time step set equal to 8 s. The two curves show that the proportional change in overall (molar) eating rate as interpellet delay increases over a 32-fold range (from 2-64 time steps, in the model, or 16-512 s in the Lucas and Timberlake data) is very similar in data and model.

Lucas and Timberlake suggest that their data on meal size imply some kind of positive feedback from each pellet (as well as

Figure 8.11 Effect of satiation value (SV) of each food portion (e.g., infusion rate, in a controlled-infusion-rate procedure, pellet size in a solid-food experiment) on total ingestion rate in the CINT model.

Figure 8.12 Effect of imposing a minimum interpellet interval (IPI) on overall eating rate. Heavy line: Simulation of CINT model. Open squares: data from Lucas & Timberlake (1988). Abscissa: IPI in seconds, except that “1” is the baseline condition of continuous availability.
regulatory negative feedback), because they found that meal size decreased at the shortest imposed interpellet interval (16 s), but further increases had no further effect. Our simulations show that the effects of IPI on overall eating rate are well explained by the CINT model, even though it lacks positive feedback. The CINT model can also explain some aspects of their meal-duration effects. It is not difficult to select SV and the time constants for the integrators (parameter \( \lambda \), see legend for Figure 8.6) so as to produce eating in meals, even at substantial IPIs. Under these conditions, meal duration tends to decrease as IPI increases, and the largest effect is at the shortest delay, as Lucas and Timberlake report.

**Meal-Intermeal Correlations.** Although there is some controversy about the details, the consensus seems to be that there is usually a weak positive correlation between meal duration and subsequent intermeal interval in free-feeding laboratory rats but negligible correlation between IMI and duration of the next meal. Even the deterministic version of the CINT model shows this asymmetry: meal-intermeal (postmeal) correlations are positive and high, intermeal (premeal) - meal correlations are positive but low. I take up this issue again in Chapter 9.

Most of the qualitative predictions of the CINT model are not very sensitive to the time constants of the integrator series. With the exception of the effects of IPI on meal pattern, the properties of the model are relatively independent of particular parameter values. Note also that although the model has five integrators, the 10 parameters are all linked as described in the legend to Figure 8.6: in this version, there is in fact only a single free parameter, \( \lambda \). The predictions of the CINT model depend much more on its cascaded structure, and the delay imposed by the initial non-satiating integrators, than on details of parameter values.

**SUMMARY and CONCLUSION**

The cascaded-integrator model seems to capture the salient features of feeding dynamics: eating in meals, the effects of long and short interruptions of feeding and regulation at all time scales. These behavioral properties all follow from the idea that feeding is inhibited by the integrated, lagged effects of ingestion. Moreover, the model is consistent with known psychobiology, which implies that feeding is inhibited by both peripheral (i.e., taste) factors and factors later in the chain of events initiated by ingestion (gastric content, nutrient uptake) and that peripheral factors generally act faster than central ones. I have not tackled the issue of incentive, what makes a given food source attractive. I look at incentive, and the role of learning in feeding and foraging, in the next chapter.

The behavior studied in reinforcement-learning experiments depends on both learning – acquired associations between behavior and environmental events; and motivation – reversible changes in the animal's internal state. These two kinds of effect are often difficult to disentangle. Because motivational effects, being reversible, are potentially simpler, we need to understand them before we can be sure of the role of learning.

Motivational processes determine how organisms allocate their time and energy among different activities as well as how the strength of a single activity varies. It is often difficult to separate these two factors: does the rat stop eating because he is no longer hungry, or because he has a sudden urge to do something else? Must we take into account the dynamics of several motivational systems in order to understand any one of them? In this chapter I have considered eating in the rat as a model system, because food is the commonest reinforcer in reinforcement learning experiments, because rat behavior resembles the behavior of many other mammals, including man, and because there is evidence that rat eating patterns can be understood in isolation from other motivational systems.

The most striking dynamic property of eating behavior is that it is regulated across several time scales: within eating bouts (meals), between meals and over very long periods (in the regulation of body weight). Regulation can be demonstrated by perturbing the system in various ways – by elevating or depressing body weight, by interrupting the normal pattern of feeding, by
spacing metered feeding episodes and by depriving the animal of food for longer periods. In every case, the pattern of feeding changes so as to more or less cancel out the effects of the perturbation. After interruption, for example, the first meal is extra long and almost compensates for lost eating opportunities. If the first meal after a period of deprivation is briefly interrupted, the amount eaten when eating resumes ensures that the total meal size is constant.

There have been numerous attempts to understand the dynamics of feeding through physiologically based models. Unfortunately, much of the necessary physiology, especially those aspects that involve the nervous system, is still poorly understood. Perhaps because of a desire to limit modeling to known physiology, early models were all more or less unsatisfactory and interest in this topic has declined since the late 1970s. Nevertheless, from a purely behavioral point of view the dynamics of eating in the rat seem to be relatively simple. I show in this chapter that the essential properties can be simulated by a biologically plausible multi-stage model. The key ideas of the model are (a) that feeding is inhibited by a satiation signal that is the sum of several internal aftereffects of ingestion; (b) that these aftereffects take place in a cascade, that can be modeled as a series of leaky integrators; and (c) that the satiation signal is taken from later steps in the cascade, hence lags after the food input. This model, which contains only negative feedback, can nevertheless explain some sensitization-like effects, eating in meals and all the major regulatory effects just described.
Chapter 9

MOTIVATION, III: INCENTIVE and SCHEDULE EFFECTS

The last chapter showed how the pattern of rat feeding seems to reflect delayed negative feedback from ingestion, the whole process modeled as a cascade of leaky integrators. This simple system obviously lacks many of the dimensions of real feeding. Nevertheless, it successfully simulates the basic regulatory dynamics: adjustment to changes in the size and spacing of meals and the effects of interruption over a wide range of time scales. In this chapter, I explore the limits of this simple regulatory system. How much of feeding and foraging can be explained by regulation alone? Where is learning involved?

I term the state of the cascaded integrator model regulatory state. It is defined by a handful of numbers (five or six in the versions discussed in Chapter 7) corresponding to the \( V \)-values of each integrator in the cascade. The input to the system, food, I defined by just a single number, termed satiation value (\( \Phi \)). This is too simple, of course: the properties of a food source or a cuisine, even a rodent cuisine, cannot be realistically represented by a single number. Food sources differ in palatability as well as satiation value – and the effects of palatability differences may depend on associative learning. Salt may be salt, to an infant as to an adult, but the taste of chili pepper, or Scotch whisky, must generally be acquired. Although I consider only a single dimension in this chapter, in a comprehensive feeding model the effects of ingested food will need to be represented by more than one number.\(^ {93} \)

Rat does not live by regulation alone. Feeding behavior is also affected by the attractiveness of a given food source. Attractiveness depends not just on the animal’s regulatory state, but also on the taste, satiation value and accessibility of food. Accessibility is another name for reinforcement schedule: factors such as where the food is located, how it is to be obtained, how much effort, of what sort, is required when, for each unit of food, etc. All the non-regulatory factors that affect the attractiveness of a particular food source – termed incentive factors – must generally be learned.

The net attractiveness of a given food source – as measured, for example, by an organism’s willingness to work to attain it – depends on incentive state as well as regulatory state. These two components, regulatory and incentive states, together comprise the organism’s motivational state with respect to a given food source.

A schematic picture of this view of the feeding system is shown as Figure 9.1. The figure shows that food has an inhibitory (satiating) effect on the regulatory system but a mixed effect (sometimes positive, sometimes negative, depending on taste and experience) on the incen-

---

\(^ {93} \) The issues discussed in this paragraph are helpfully reviewed in a chapter on modeling by Toates and Evans (1987). See also a good discussion in Weingarten (1985).
Adaptive Dynamics

Staddon

tive system. The outputs of the regulatory and incentive systems are combined in ways to be discussed in a moment to produce some level of operant behavior, i.e., behavior motivated by the search for food or some other reinforcer. The effectiveness of food-motivated operant behavior (foraging) depends on the environment in which it occurs. In experimental studies, the constraining effects of environment are represented by the reinforcement schedule, which determines how much of what response, when, shall procure a portion of food. The food obtained then has feedback effects on regulation and incentive, as just described. The operant behavior itself may be an input to the incentive system and will presumably have generally negative effects (the more you have to work, the less attractive the food source), although these may be rather small under some conditions, as we saw in the last chapter.

Incentive, particularly if it involves associative learning, involves processes more complex than regulation. The principle of parsimony enjoins that we begin as simply as possible. My first task, therefore, is to see how much of feeding and foraging behavior can be explained without resorting to learning, i.e., by regulatory processes alone, making the simplest possible assumptions about the incentive system. When regulatory ingenuity begins to fail, something more complicated may need to be introduced. So, in this chapter I look at situations on the borderline between regulation and associative learning and try to understand which effects are on which side of the border. I will conclude that regulation, plus the most minimal assumptions about incentive, can account for a surprising range of phenomena. There is less need for learning than is usually assumed.

Model Uniqueness

It will strike many readers that the feeding model I developed in the last chapter (Figure 8.5, Equations 8.7-8.10) might be modified in several ways. Since I make no attempt to link elements in the model directly to physiology, parsimony is essential. How unique is the model discussed in Chapter 8? Is it the simplest possible model that can explain the behavioral data? Before we get to applications of the model to new data, it may be helpful to say something about alternative versions.

Here are some obvious modifications. As Figures 8.4 and 8.5 illustrate, in the original model (cascaded integrator - simple: CINT-S), the input to integrator \( i \) is just the output of integrator \( i-1 \). A very similar model is illustrated in Figure 9.2, but in this case the input to integrator \( i \) is not the output of integrator \( i-1 \) but the difference between the output of integrator \( i-1 \) and the state of integrator \( i \). The equations for this model are:

\[
x(t) = \phi(t) \text{ if } V_S < \theta; \ 0 \text{ otherwise}, \quad (9.1)
\]

\[
V_i(t+1) = a_i V_i(t) + (1-a_i)x(t), \quad (9.2)
\]

\[
V_i(t+1) = a_i V_i(t) + (1-a_i)[V_{i-1}(t)-V_i(t)]. \quad 0 < i < N, \quad (9.3)
\]

These equations, for model CINT-D (D for difference), are identical to those for CINT-S, except for the last term in the third equation.

My colleagues and I have also looked at a third version, CINT-F, in which the decrement in the \( V \)-value of the last integrator is fixed (representing a constant metabolic rate), rather than proportional to \( V \).
The version of CINT-S discussed in the previous chapter has five stages. How many are correct? Feedback was taken only from the last two stages: is this right, or could all units contribute without changing the major characteristics of the model? We have looked at a five-stage version of the model where the feedback comes from the last two, the first two, or all five stages. When feedback comes from either all, or the last two stages, the model regulates in response to different “bite” sizes, i.e., the total intake is approximately constant when SV size varies over at least an eightfold range. This corresponds to the universal empirical finding that rats will maintain a more or less constant caloric intake on diets with food of different caloric densities. The model, like real animals, doesn’t regulate perfectly: the richer the diet, the more calories are ingested, but not in proportion: doubling the SV of each food portion produces only a small percentage increase in total intake (see Figures 8.7, 8.10).

The biphasic impulse response (i.e., delayed response) of models in which feedback comes only from the later stages (see Figure 8.6) is essential to eating in meals. A three-stage model, with feedback from the last one or two stages also regulates and yields eating in meals, but the eating pattern is less variable than a 5-stage model. So which version is best? The fact is that all three CINT models perform similarly: the details don’t seem to matter very much. The critical thing seems to be the biphasic impulse response with a slowly declining “tail” (Figure 8.6, Figure 9.19, below). A function like this implies that there are several integrators in cascade and that the satiation signal depends more on later elements in the chain than earlier ones. In the remainder of the chapter I will concentrate on a 3-stage version of CINT-S, with feedback from the last stage, because it seems to be the simplest version that can simulate the regulatory properties of feeding and foraging.

Circadian Effects

Rats feeding freely prefer to eat during the dark phase of a 12:12 light-dark daily cycle (Siegel, 1961). Meals in the dark are large and closely spaced, in the light small and far apart. This is not an obligatory pattern. As we noted in the last chapter, if more food is available in the light than in the dark, rats adjust their feeding pattern accordingly. The circadian pattern is easily abolished if access to food is even slightly restricted (Levitsky, 1974). How should this diurnal variation be accommodated in a regulatory model? In the following discussion, I deal with diurnal variation by assuming a fixed-proportion step change in threshold, \( \theta \), between the light and dark periods of the 12:12 cycle.

Incentive Model

The animal’s tendency to eat (“reflex strength”) is necessarily just a single number. In the model, strength is equal to threshold - satiation value: \( H = \theta - V_s \), where \( V_s = \sum w_i V_i \), the weighted sum of “state” variables. \( \theta - V_s \) can be thought of as a measure of “hunger,” \( H \), although this identity should not be taken too literally.

But reflex strength (vigor of eating, or an animal’s tendency to work for food), depends on incentive as well as hunger (Figure 9.1). Some foods are more attractive than others: rats will generally press a lever more vigorously to get more concentrated sugar solutions, and will lick them faster, for example (cf. Figures 7.5 & 7.6). Taste and “richness” (portion size, caloric density – satiation value) both affect the attractiveness of a food source. Some assumption about incentive is necessary, no matter how simple the experiment.

Incentive always has some learned components. Even if a given food is “innately” attractive, the animal must still learn where it is to be found, learn to associate it with predictive stimuli, and so on. But, since my aim is to see how much can be explained by regulatory state alone, I will assume initially that incentive value is proportional to satiation value (food portion size or caloric concentration), without making any assumptions about how SV is assessed, whether it is learned or innate, etc. I will not deal with dimensions of food other than satiation value.
In the feeding experiments discussed in the last two chapters, the only behavior required of the animal was eating. (I assume that going over to the feeder is a costless, overlearned activity, and neglect its dependence on learning for the moment.) But in the experiments to be discussed now, an “arbitrary” response (sometimes many responses) like lever pressing is required for access to food. A lever-press requirement may have a direct effect on incentive. If many responses are required for each small food portion, the incentive value of the food source should be low; if few responses are required, or the food amount is large, incentive value should be high. There may also be long-term regulatory effects as a consequence of the energy expenditure required by large lever-press requirements. In fact, we’ll see in a moment that the usual lever-press requirement is treated as relatively costless. It seems to exert its negative effects indirectly, through the time delay imposed by satisfying the requirement. But any immediate effect of a lever-press requirement independently of its effect on the time between food deliveries will presumably involve associative mechanisms. There are certainly situations where a lever-press or other operant requirement can have effects on incentive, and thus involve associative processes, as well as on hunger (which involves only regulatory processes), but these situations may be rarer than usually assumed.

Gating. Presumably the animal’s net tendency to engage in lever pressing, the reflex strength of lever pressing, depends in some way on both regulatory state and incentive. How are the strength values corresponding to these two states to be combined? The two simple combination rules are: additive and multiplicative (also termed “gating”). Of these two, only the gating option is plausible, because a food source that is inaccessible (i.e., impossibly high cost, hence zero incentive value) should never be responded to, no matter what the organism’s regulatory state. In what follows I will assume that the strength of the animal’s tendency to press lever $i$ for food, $B_i$, (the reflex strength of lever pressing), is proportional to the product of its internal regulatory strength ($H$) and the net incentive of that food source (call it $I_i$):

$$B_i = kHI_i,$$

(9.4)

where $k$ is a constant.

The way this incentive assumption works is illustrated in Figure 9.3. The figure shows reflex strength for a 5-stage model as it ingests foods with satiation values of 6 (heavy line) or 3 (light line), beginning with a zero initial condition (extreme deprivation, $V_i(0) = 0$ for all $i$). The regulatory state in both cases is the same, but the initial reflex strengths are in 2:1 ratio, reflecting the 2:1 ratio of satiation (incentive) values. The SV-6 curve declines more rapidly than the SV-3 curve, reflecting the greater satiating effect of the larger food portion. The general form of

---

94 Lever pressing is not a reflex in the traditional sense of the term. But I retain the term “reflex” for any response system that has a strength measure.

95 The gating assumption seems to have been first proposed by Yale theorist Clark Hull (1884-1952). Hull founded a behavioristic school of “rat psychology” that was dominant during the 1930s through the early 1950s. He attempted to model static data from group experiments using trial-by-trial procedures, rather than the individual, real-time data needed for dynamic analysis. Hull’s theory has been largely abandoned, but he did lay out the problem to be solved and several components of his scheme, including the multiplicative relation between hunger (Hull’s “drive,” D) and incentive as determinants of reflex strength (“reaction potential”), are retained by contemporary models for animal learning (Hull, 1943, 1952; Osgood, 1953).
the two curves resembles results from meal-duration experiments with different sucrose concentrations (compare Figure 9.3 with Figure 7.6). Any CINT-type model with the properties we have described will behave in this way.

Given these minimal assumptions about incentive state, we are now in a position to see how well the CINT models can account for experiments on foraging.

THE COLLIER EXPERIMENTS

George Collier and his associates have for several decades been carrying out systematic behavioral experiments on feeding and foraging in rats. These studies have led Collier to emphasize the importance of what he terms *feeding strategies*. He acknowledges the importance of regulatory physiology (Collier, 1986), but contrasts his own ecological and economic view with what he calls the *standard homeostatic model*:

The standard model is one in which deprivation results in depletion and hunger, and surfeit results in repletion and satiety. Feeding behaviors are presumed to begin and end *in response to* these deficits and surfeits...This model is embedded in the concept of *homeostasis* (Collier & Johnson, 1990, p. 771).

![Collier Procedure](image)

**Figure 9.4** Collier’s experimental procedure. The figure shows the procedure used in most of the single-choice feeding experiments by Collier and his associates. Operant responses (usually, lever-pressing) are of two types: “procurement” or “access” responses and “consumption” responses. Procurement responses (N_p) are made for the opportunity to engage in consumption responses (N_c). Consumption responses are made for the opportunity to eat. (This is called a *chain* schedule, in operant terminology, but Collier uses his own terms.) The procedure works like this: When the animal has ceased to make a “consumption response” (defined below) for experimenter-set time D (D is usually 10 min), further consumption responses are ineffective and the animal must make responses on the *procurement* or *access* lever to get further food. After he has made P procurement responses (P = 11 in the figure, but the usual requirement is much larger), a “meal opportunity” is offered. Now the animal must work on the *consumption* lever to get actual access to the food. After C “consumption” responses (C = 2 in the figure; usually C << P, consumption cost is much less than procurement cost), a fixed amount of food is delivered. The animal can get additional food portions by paying the consumption cost again and this possibility remains available indefinitely – until the animal again ceases to make consumption responses for time D, at which point the procurement cost must be paid again and the cycle repeats. In some experiments, the only consumption response required is eating itself, detected by a photocell arrangement. See text for other details.

Collier and his coworkers have produced many experimental results that appear to contradict the standard model:

When an animal can detect the structure of its habitat, its feeding behavior will tend to reduce the feeding cost/benefit ratio in apparent anticipation of both its nutritional requirements and the environmental conditions. This implies a system controlling when, where, and on what source the animal feeds, and how much and how fast it eats on any occasion, that is *quite different from the traditional homeostatic model*....[H]unger and satiety are not necessarily correlated with deficits and surfeits but rather with the patterns of intake required by different feeding strategies (Collier & Johnson, 1990, pp. 775-6; [emphasis added]).
Collier contends that the apparently anticipatory feeding patterns in his data depend upon associative learning. I contend that the role of associative learning in feeding regulation may be much less than is often assumed. In this section, I take up the challenge of showing that the major properties of the feeding data that Collier has interpreted in terms of learned anticipation can in fact be explained by the homeostatic model introduced in the previous chapter.

**Procedures.** Collier uses procedures that differ in critical ways from “standard” operant conditioning: the animals are not food deprived, other than through limitations on their ability to get food imposed by the reinforcement schedule, and “reinforcement” is not a fixed amount of food, but a “meal” of indefinite duration. Collier’s aim is to mimic natural foraging in which isolated meals (termed “consumption” by Collier) are separated by periods of searching for the next meal (termed “procurement”). Experimental “sessions” are continuous throughout the 24-hour day, except for a brief daily maintenance period. This is sometimes called a *closed economy* (see Chapters 2 and 3), to be distinguished from the usual brief-session-followed-by-free-food-top-up *open economy* in standard conditioning experiments.

Collier’s basic procedure is diagrammed in Figure 9.4. Look over the figure carefully, because these procedural details are very important to the behavior that animals show in these experiments. The basic idea is that the animal must make a large number of responses (usually lever pressing, for rat subjects) for the opportunity to eat (this requirement is termed “procurement” or “access” cost: $P_c$); but once food is available, it remains available (at small, or no, “consumption” cost: $C_c$) until the animal ceases to eat or make the consumption response for a substantial period (usually 10 min), at which point it must again “pay” the access cost.

Collier’s procedure permits him to manipulate procurement cost and consumption cost either chronically (steady-state) or more frequently and look at the effect on eating. I’ll look at the pattern of eating and operant responding after indefinite exposure to a given procurement or consumption cost (steady state). I’ll also look at one experiment that studied the pattern of eating and responding when procurement and consumption cost change more frequently. How well can the CINT model simulate these effects?
Procurement Cost

If procurement cost is increased (and consumption cost is low) the effect is jointly on meal frequency and meal size. Meal frequency decreases, which means that the animal needs to pay the procurement cost less often. Meal size increases, so that the animal does not suffer too much reduction in its total food intake. Typical steady-state results are shown in Figure 9.5 for an individual rat (Collier, Hirsch, & Hamlin, 1972), in a closed economy with food available according to the procedure diagrammed in Figure 9.4. Consumption cost was zero in this experiment: food was made unavailable 10 minutes after the last eating episode, as measured by a photocell arrangement. The open squares (righthand y-axis) show per-diem frequency of “meals” (defined according to the 10-min criterion) at different procurement costs, from 1 (continuous reinforcement; CRF) through 5120 (i.e., a fixed-ratio, FR, 5210 schedule). The open triangles (lefthand y-axis) show mean meal duration in minutes.

As you can see, meal frequency decreases as procurement cost increases, but meal duration goes up. Figure 9.6 shows that these two effects more or less compensate for one another: as meal frequency decreases, meal duration increases so that total intake is more or less constant despite a 5000-fold increase in the response-cost of each meal. Only when procurement cost is greater than about 300 responses does total intake begin to decline. Thus, the animal behaves in a highly adaptive fashion: decreasing meal frequency as procurement cost increases (so the cost need be paid less and less often), but compensating for the resulting drop in eating rate by large, largely costless, increases in meal size. Very similar effects have been found in cats and guinea pigs (Kanarek, 1975; Hirsch & Collier, 1974; see review in Collier, Hirsch & Kanarek, 1977). How might these results be explained homeostatically? Quite simply, it turns out. The argument is as follows. As procurement cost (fixed-ratio size) is increased over the very large range used in these experiments, the time between “meals” must also increase beyond its previous value. Why? Because the rat does not increase its lever-press rate significantly at longer ratios, so it takes more time to finish the requirement at higher ratios. The question for the CINT models is whether this forced increase in intermeal interval is by itself sufficient to explain the effects of procurement cost.

If you recall the effects of interruption discussed earlier, you are likely to answer “maybe,” because following a forced interruption of feeding, the first meal is always extra-large – and this effect is predicted by the CINT models. This is also what Collier’s experiments show: when the procurement cost is high, so that intermeal interval is larger than it would be under unconstrained conditions, meal size is also large. Recently, Mathis, Johnson and Collier (1995) have shown that procurement time and procurement cost are roughly equivalent: “The functions relating meal frequency and size to the procurement interval were of the same shape as those seen when cost is the completion of the bar-press requirement...(p. 295)” The decline in meal frequency appears to be an adaptive response to increased procurement cost. But it may also be an incidental byproduct of the time taken to complete large procurement-cost ratios.

It’s worth noting that there is no contradiction between these two interpretations: The change in behavior caused by increasing procurement ratios can be both adaptive – as Collier asserts – and the outcome of a simple regulatory process.

Simulations96

In this section I present detailed simulations of data from four studies. We have chosen almost the simplest possible model with the necessary biphasic impulse response: three stages (two would be the minimum possible), with feedback only from the last stage (in Figure 8.5, \(w_1\), and \(w_2\) equal to zero and \(w_3 = 1\)). The parameters were chosen to be the best single set to give

96 These simulations were done in collaboration with Silvano Zanutto (see also Staddon & Zanutto, 1998; Zanutto & Staddon, 2007).
good predictions for all the experimental foraging data we discuss in the rest of the chapter. No explicit “best-fit” procedure was used. Our objective was to see if any version of the 3-stage model could provide a qualitatively adequate fit to a wide range of experimental “foraging” data (remember, the essential feature of this model is the biphasic impulse response, not the number of integrator stages). Model details are given in Box 9.1.

I begin with two simple experiments already described in Chapters 7 & 8. The first, by Levitsky et al. (1976), just involved depriving rats of food for varying periods and looking at the effect on their eating pattern. The second, by Lucas and Timberlake (1988), imposed a minimum time delay between successive food pellets. We then attempt to model the Collier et al. (1972) study and finally look at an elaborate study by Morato, Johnson and Collier (1995) in which the procurement and consumption requirements changed in different patterns.

**Box 9.1 Assumptions of the Model**

We used the same 3-stage model to simulate experimental data from four studies: Collier Hirsch & Hamlin (1972, Expts. 1 & 2), Levitsky, Faust & Glassman (1976) Lucas & Timberlake (1988) & Morato, Johnson & Collier (1995). We made the following assumptions:

- The model iterates second by second (1 time step = 1 s).
- The model assumes that the threshold, $\theta$, is lower during the day ($\theta = .09$) than during the night ($\theta = .13$), so that the model usually eats more at night (12:12 cycle).
- It takes time to consume a pellet. We assume 8 s of eating, followed by a refractory period of 9 s before the next activity.
- The model lever-presses at a fixed rate: 2/s, so long as $V_s < \theta$, it is not eating and no refractory period is in effect.
- Meal begins with the first pellet; when $V_s \geq \theta$ eating ceases. Meal duration, $t_M$, is defined as time between the first pellet and 60 s with no eating, $t_m$, less 60 s: $t_M = t_m - 60$.
- When an experiment starts, $V_1$ and $V_2$ are zero; $V_3 = 0.84$ to reflect the previous day’s feeding.

The equations of the model are as follows (see equations 8.8-8.10 in Chapter 8):

\[
x(t) = \phi(t) \text{ if } V_5 < \theta \text{ and food is available; } 0 \text{ otherwise, (9.5)}
\]

\[
V_i(t+1) = a_i V_i(t) + (1-a_i)X(t), \quad (9.6)
\]

\[
V_i(t+1) = a_i V_i(t) + b(1-a_i)V_{i-1}(t), \quad 0 < i < N, \quad (9.7)
\]

Parameter values for all the simulations were set as follows: $a_1 = \lambda_1$; $a_i = a_{i-1} + \lambda_2(1-a_{i-1})$; $b = 1.3$. In these simulations, $V_3$ was the $V$-value for the last stage (i.e., $V_5 = V_3$), $\lambda_1 = \lambda_2 = 0.977$, and $\phi = 2$. Thresholds: $\theta_{\text{dark}} = .13$, $\theta_{\text{light}} = .09$.

**Figure 9.7** shows simulations of this model for the experiment by Levitsky et al. discussed in Chapter 7. In this experiment, groups of free-feeding rats were deprived of food for 1, 2, 3 or 4 days. The figure shows cumulative amount eaten, per rat, in g, across days before and after the interruption. In every case, the group eats more just after the interruption, but the cumulative curve fails to rejoin the pre-interruption curve; the shortfall is substantial for interruptions.
greater than 24 hours. The steady-state eating rate is approximately the same after interruption as before. The model (solid line) matches these properties closely, although the rate of eating is a little higher than the rate of the rats used in this study (remember that parameter values were set for a best fit to all the data in four different studies).

Figure 9.7 compares the prediction of the Box 9.1 model (solid line) with data (open squares) from the Lucas and Timberlake (1988) experiment described in the last chapter. In this experiment, undeprived rats in a closed economy were constrained to receive pellets (produced by nosing into a trough) at minimum interpellet intervals, ranging from 10 s to 512 s. The fit to data is not as good as with the five-stage model shown in Figure 8.12, but the general pattern is the same.

Figure 9.9 shows predictions of the same 3-stage CINT model for Experiment 1 of Collier et al. (1972). In this experiment, undeprived rats in a closed economy had to gain access to a “meal” by making anywhere from 1 to 5120 “procurement” lever presses. Consumption cost was zero in this experiment. Each cost requirement remained in force for five days. As you can see, the CINT model predicts the meal-size, meal-duration and meal-frequency data reasonably well; the model regulates total intake rather too well – better than the rat at high ratio values. Nevertheless, an individual rat that behaved like the CINT model would not be
significantly different from the other rats in Collier et al.'s study (other than showing less variability). The model is within the range of the animals in this experiment.

Thus, the main effect of procurement cost on meal size is consistent with the same process that we used in the last chapter to explain eating in meals and the first-meal effect of feeding interruption. The decrease in meal frequency with increasing procurement cost, and the maintenance throughout of close-to-control levels of daily ingestion, is also predicted by the CINT model.

There are some apparently contradictory features of the data, however. As procurement cost is varied, correlations between meal size or duration and subsequent intermeal interval are high in the CINT model: long meals are followed by long intermeals, short meals by short intermeal. But within most of Collier's experiments these correlations are low. The absence of meal-intermeal correlations in the data is one of Collier's main arguments against a regulatory interpretation. I return to this issue in a moment.

Figure 9.10 shows predictions of the model for Experiment 2 of Collier et al. (1972), in which consumption cost was varied and procurement cost was set to zero. The match of model to data about as good as good as in Experiment 1. Meal size varies little with consumption cost (upper-left panel) in model and data, but the variation takes a different form. On the other hand, meal duration is directly related to consumption cost in both model and data (upper-right panel). Daily intake declines slightly in both model and data as consumption cost increases (bottom-left panel), but the model regulates a bit better. Finally, model and data show bitonic, but opposite, variation in meal frequency with consumption cost (bot-
But notice that the low-consumption-cost meal frequency in this experiment is much higher (~27/day) than is typical (compare with Figure 9.9, where the free-feeding meal frequency is more typical: about 9/day). This discrepancy makes it hard to interpret the disparity between meal-frequency data and prediction at the low-consumption-cost end in Figure 9.10.

**The Morato, Johnson & Collier (1995) Study.** The next six figures compare predictions of the CINT model with data from an elaborate experiment by Morato, Johnson & Collier (1995) in which conditions were varied more or less frequently. The experiment held consumption cost constant at 10, and looked at three levels of procurement cost, 10, 200 or 400. Conditions were alternated daily (e.g., 1x10, 1x200) or in more complex patterns: 5x10, 5x200 (five days at 10 followed by five days at 200 and then repeat); 10x10, 10x200, etc. and comparably for the 400 cost. Each alternation condition continued for 40 days. The complete set of conditions is shown in the figures. Average data are presented in the published paper, but we have obtained individual data from George Collier, some of which are shown in the figures.

**Figure 9.11** The effect of procurement cost (ratio value) in the 10x10, 10x200 condition (10 days at a procurement cost of 10 alternating with 10 days at a procurement cost of 200) in the experiment of Morato, Johnson & Collier (1995): simulation (solid triangles) and average data (solid lines, filled circles). Open symbols show data from two individual rats. The three panels show meal frequency, meal size and daily intake.

**Figure 9.12** The effect of procurement cost (ratio value) in the 5x10, 5x200 condition in the experiment of Morato, Johnson & Collier (1995): simulation (solid triangles) and average data (solid lines, filled circles). Open symbols show data from two individual rats. The three panels show meal frequency, meal size and daily intake.
two rats (open symbols). In this figure, and in those to follow, the model captures the pattern of the data but shows effects that are smaller. Predictions of meal frequency changes are better than predictions of change in meal size. Look at the left-hand panel of Figure 9.11, which shows meal frequency. During the

![Figure 9.13](image)

**Figure 9.13** The effect of procurement cost (ratio value) in the 1x10,1x200 condition in the experiment of Morato, Johnson & Collier (1995): simulation (solid triangles) and average data (solid lines, filled circles). Open symbols show data from two individual rats. The three panels show meal frequency, meal size and daily intake. The first 10 days, cost is high, during the next 10, low, and so on. Both model and data show that meal frequency is high when cost is low and low when cost is high. Conversely, in the center panel, meal size is large when cost is high, small when cost is low. The right-hand panel shows that daily intake varies relatively little across conditions, more in the data (and at a lower absolute value) than in the model. The open symbols show that the individual rats varied considerably about the mean. **Figure 9.12** shows five-day alternation and both model and average data repeat this general pattern.

**Figure 9.14** The effect of procurement cost (ratio value) in the 1x10,1x400 condition in the experiment of Morato, Johnson & Collier (1995): simulation (solid triangles) and average data (solid lines, filled circles). Open symbols show data from two individual rats. The three panels show meal frequency, meal size and daily intake.
variation in meal size and frequency is similar in average data and model, but daily intake regulation (right panel) is better in the model.

The next three figures show similar data, and model predictions, for the 10,400 conditions (Morato et al., Figure 2). **Figure 9.14** shows meal frequency and size and daily intake data, and model predictions, for the daily alternation condition. The model predictions are similar to those for the 200-condition, but the model shows less daily variation in meal frequency (left) and meal size (center) than the average data and much less variation in intake (right).

![Figure 9.14](image)

**Figure 9.14** The effect of procurement cost (ratio value) in the 1x10,2x400 condition in the experiment of Morato, Johnson & Collier (1995); simulation (solid triangles) and average data (solid lines, filled circles). Open symbols show data from two individual rats. The three panels show meal frequency, meal size and daily intake.

**Figure 9.15** shows data for the 1x10,2x400 condition and the same basic pattern is repeated: frequent, small meals on low-cost days, few, large meals on high-cost days. The model regulates much better than the rats, but in this case as in all the others, the pattern of variation in intake – less on high-cost days than on low-cost days – is the same as the data. The overall pattern of results and simulations in the 1x10,4x400 condition (**Figure 9.16**) is the same.

We conclude that the regulatory model can reproduce the pattern of results in a wide range of feeding experiments, including some operant foraging studies. The predictions seem to be best in those situations where the animals can maintain their body weights well (Levitsky & Glassman, Timberlake & Lucas, Collier, et al., Experiment 1) less good where real rats fail to regulate (Collier et al, Experiment 2, Morato et al.).

![Figure 9.15](image)

![Figure 9.16](image)
Meal-Intermeal Correlations

Many studies of free feeding in rats have found correlations between the length (or size) of a meal and the subsequent intermeal interval (Clifton, 1987; Le Magnen & Tallon, 1966; Levitsky, 1974; Panksepp, 1973, 1976). We summarized the rather conflicting data in Chapter 7: meal-intermeal (M-IMI) correlations (where they have been found) are moderately positive (r = 0.1 - 0.7); meal-premeal (IMI-M) correlations are small or even slightly negative. These correlations are affected by several variables: (a) Interruption: if feeding is interrupted for 24 hours or more, the first post-fast meal is extra long. Under these conditions, the IMI-M correlation is obviously positive, not zero. (b) Effort: M-IMI correlation increases slightly as more effort (e.g., holding down a bar for a specified time) is required to gain access to food, and M-IMI correlation goes from zero to small positive values (Levitsky, 1974). (c) Meal size: Manipulations that increase meal size (measured in calories) seem to increase M-IMI correlations. Levitsky (1974) showed that adding fat to the diet or adding a response requirement, both of which increase meal size, also caused an increase in M-IMI correlations. Adding cellulose or quinine to the diet, both of which reduce meal size, also reduced M-IMI correlations. Studies that have found M-IMI correlations seem to use procedures that ensure large meal sizes, whereas meal sizes are small in studies that have failed to find correlations.

To what extent can this pattern of results be duplicated by the CINT model? It is particularly important to understand the conditions under which the CINT model may fail to show correlations, because correlations between (for example) intermeal duration and meal size are often assumed to be reliable if not infallible signatures of a regulatory process, so that lack of correlation has been used to rule out regulation. The absence of these correlations in some feeding data one reason that Collier and others have downplayed feedback models in favor of more complex accounts of behavior in experimental foraging paradigms.
Simulations. The CINT model is strictly deterministic. In the steady state, in the absence of a perturbation, meal and intermeal durations are fixed, so there will be no correlation between M and IMI or the converse. In real data, M and IMI do vary. If the variation (assuming the model to be basically true) reflects variation in time scale (or, equivalently, in rate parameters) the model also predicts no M-IMI or IMI-M correlations under undisturbed conditions. The assumption of circadian variation in threshold does constitute a perturbation and does cause some variation in M and IMI in the model and does lead to correlations, which are summarized in Figure 9.17. The figure shows the effect of SV magnitude (Φ) and procurement cost, on the size of the M-IMI and IMI-M correlations induced by diurnal changes in θ. As in the data, M-IMI correlations tend to be positive and IMI-M correlations zero or negative. The positive M-IMI correlation tends to increase as Φ increases, as in Levitsky’s data. The smaller negative IMI-M correlation tends to decrease. Procurement cost has no consistent effect on the M-IMI correlation, but tends to drive the negative IMI-M correlation more negative. The general pattern in Figure 9.17 – relatively high M-IMI correlations, low or negative IMI-M correlations – is also the commonest pattern in empirical studies.

The major experimental implication of this analysis is that for a free-feeding rat under uniform 24-hour conditions, correlations are expected to be small. But anything that causes a change in threshold, such as the introduction of a competing activity like wheel running, should immediately produce M-IMI and IMI-M correlations of the type illustrated in Figure 9.17. Experiments of this sort do not appear to have been done.

CONCLUSION

Chapters 8 and 9 showed that: (a) hunger-motivated operant behavior is driven by both regulatory and incentive factors, which combine nonlinearly (a multiplicative rule seems to work). (b) The qualitative pattern of both feeding constraint and perturbation data (Levitsky & Glassman, Timberlake & Lucas) and experimental foraging data (Collier et al.) can all be simulated by a simple 3-stage cascade model with the same fixed parameters. (c) The model also simulates the general features of meal-intermeal correlations and of manipulations that affect meal-intermeal correlations, but (d) The model predicts changes from condition to condition that are often quantitatively too small in the Morato et al. experiment (Figures 9.11-9.16). Model and data differ in absolute values for daily intake, meal frequency and meal in several comparisons.

What are we to make of the model’s apparent failures? First, it is necessary to get a feel for just how large those failures are. Figure 9.18 shows predictions of a 3-stage model with slightly different parameters (Model 1; the model discussed earlier is Model 2) for the 1x10, 1x200 condition of the Morato et al. study (compare with Figure 9.13). Model 1 does better than Model 2 in predicting changes from condition to condition, but little better in predicting absolute values. The meal-frequency prediction (filled triangles) is about the same, the meal-size prediction better and the intake prediction a bit worse (too high rather than too low). The data are quite variable: the mean is not always representative of the individuals. For example, while the average (filled circles) shows daily alternation in intake, some of the individual animals (e.g., # 4, open triangles) do not (both simulations show little daily intake variation).

The difference between the two models is in the impulse responses, shown in Figure 9.19 and some other parameters, described in the figure legend. Both models have the same general form: ingested food has no effect on satiation for a few minutes and its satiating effect does not peak until about 2 hours after feeding. It appears that even the complex data from the Morato and Collier experiments are consistent with the type of biphasic satiation process shown in Figure 9.19. Given some variability across experiments and a few anomalous results such as the

97 Assuming that the threshold in our model reflects competition with other motivational systems, so that adding a competing activity in effect lowers the threshold for satiation (raises the threshold for hunger).
meal-frequency data in Figure 9.10, our Model 2 does about as well as can be expected. It seems fair to conclude that the major pattern in all these feeding and foraging results is the outcome of a lagged regulatory process.

What about learning? Well, there surely is some role for learning, if only in acquisition of the operant response in the experimental context, and in timing meal duration (cf. Davis & Smith, 1990). But we need to take seriously the idea that the basic patterns of behavior in these studies are determined by a simple regulatory process, with learning mechanisms coming in as a later, secondary effect.

**Figure 9.19** Impulse responses (i.e., $V_s$ following a single piece of food, starting from zero initial conditions) for the model used to simulate data in Figures 7.7 through 7.17 (Model 2) and another version, Model 1. The parameters for Model 2 are given in Figure 8.6. The parameters for Model 1 are $\lambda_1 = \lambda_2 = .978$; eating only 12 hours out of 24, $\theta = .13$; food amount, $\Phi = 1.8$; eating time, 7 s, post-eating refractory period, 10 s; other details as in Model 2. Notice the logarithmic time scale.
Chapter 10

ASSIGNMENT OF CREDIT

Success in life depends on connecting cause and effect. Animals generally act to promote good ends and avoid bad, which means learning what causes what. When something of consequence happens, the animal must ask (a) Did I cause it, and if so how? This is the operant assignment of credit problem. And (b) Can I anticipate it next time? In other words, was this good or bad thing preceded by some reliable signal? This is the Pavlovian assignment of credit problem.

The term assignment of credit arose in early artificial intelligence (AI) discussions about how a learning system can decide which event – or activity, or change in its own internal state – is the actual cause of something that has happened to it (Minsky, 1961). Although they don’t use the AI term, several accounts of operant and classical conditioning (Dickinson, 1980; Revusky, 1977; Staddon, 1983) argue in effect that the laws of conditioning have evolved to optimize credit assignment in natural environments. Because the function of life is not knowledge but action, operant assignment of credit is, from an evolutionary point of view, the more fundamental. It is also simpler. The main part of this chapter is about operant assignment of credit. I summarize some of the problems posed by Pavlovian assignment of credit at the end of the chapter.

OPERANT ASSIGNMENT OF CREDIT

Until quite recently most learning models took the assignment-of-credit problem for granted. They treated all learning as supervised learning. In supervised learning a “teacher” (reinforcement) is assumed to present explicit error information to the learning system. For example, in the old Bush-Mosteller (1951) stochastic learning model and its many descendants reinforcement is assumed to selectively strengthen the “reinforced response” without any explicit discussion of how the organism – the learning system – knows what that response is.

With few exceptions (e.g., Sutton, 1984) assignment of credit was satisfactorily solved for most learning theorists by the intuitively appealing but unformalized mechanism of temporal contiguity. The informal theory amounts to little more than the assertion that activities are “strengthened” by contiguity with reinforcement, and the more closely the reinforcement follows the activity, the greater the strengthening.

Any informal theory creates its own anomalies, and the vague version of response-reinforcer contiguity learning is no exception. The crisis in the late 1960s and early 1970s caused by so-called “biological constraints on learning,” in the form of apparent exceptions to the reinforcement principle such as autoshaping, superstition and instinctive drift (Cf. Breland & Breland, 1961; Seligman & Hager, 1972; Staddon & Simmelhag, 1971), dealt “general learning theory” a blow from which it is yet to recover.

Among other things, these effects showed that “operant” activity can occur even though it has no effect on reinforcement (superstitious behavior); and the activity that does produce reinforcement may not be the one actually strengthened (instinctive drift). These facts had not been squared with the standard contiguity account of operant reinforcement as recently as 1988 (Gardner & Gardner, 1988). In the rest of this section I describe a real-time model (slightly modified from Staddon & Zhang, 1991) that offers one kind of solution to the assignment-of-credit problem in operant conditioning. The model also shows how to reconcile at least some of the exceptions to the informal idea of contiguity learning.

Reinforcement: Standard Properties

Certain kinds of events, food, access to a mate, freedom from restraint, electric shock – reinforcers and punishers – have selective effects if their occurrence (or nonoccurrence, in the case of electric shock) is made to depend on the occurrence of some identifiable activity. Under
most conditions – most responses, most kinds of response-reinforcer dependency – the effective activity is facilitated relative to others. Many years of research have revealed a number of what might be called standard properties of operant reinforcement, as well as a number of anomalous properties. The standard properties are selection, and the effects of delay and contingency. The anomalous properties are superstitious behavior, instinctive drift and autoshaping.

Selection. Reinforcement is defined by its selective effect: when the occurrence of a reinforcer is made to depend on the occurrence of a particular response, the response probability should increase, and when the reinforcer is no longer presented, or is presented independently of responding, response probability should decline. Selection must be reversible between at least one pair of activities: capable of strengthening reinforced activity A at the expense of activity B that is not reinforced, and vice versa.

Delay. In the usual operant conditioning situation, the reinforcer (food, typically) follows the operant response immediately. But if there is a delay, the reinforcer is much less effective. The delay can be imposed in two ways: as a causal delay or as a mandatory delay. A causal delay is when a response at time \( t \) produces a reinforcer at time \( t+D \), with no further constraint. A causal delay permits additional responses to occur at intermediate times, \( t+k \), where \( k<D \), without affecting the delivery of the reinforcer at time \( t+D \). The problem with this method is that the actual, experienced delay between the response and the reinforcer is under the control of the animal: if it responds quickly, the actual delay may average much less than the programmed delay, \( D \).

The alternative is a mandatory delay, which is enforced by resetting the delay timer with each response. This method ensures that the response-reinforcer delay will always be at least \( D \), but at the cost of providing the animal with a temporal cue: the fixed time between response and reinforcer. It can now learn not to respond not because its “response strength” has been directly weakened by the response-reinforcer delay, but because it has learned the time gap between responding and reinforcement, and so waits after each response to see if reinforcement is forthcoming. Moreover, this schedule actually punishes high response rates by deferring reinforcement until the response rate falls below the prescribed delay.

There is no experimental solution to this dilemma. There is no way of measuring the effects of “pure” delay. All we can do is observe the behavior under each delay condition and see what kinds of model are compatible with it. The data are clear: under both kinds of delay, responding is weakened. For example, if we train animals under different values of delay, \( D \), we find that the longer the delay, the slower the acquisition and the lower the final response rate (e.g., Dickinson, Watt, & Griffiths, 1992). Not surprisingly, the mandatory delay method has a larger effect than the causal delay method, even if the obtained rate of reinforcement is matched between the two conditions.

Contingency refers to a simple kind of predictiveness. For example, suppose that the probability a response will produce immediate reinforcement is \( q = .3 \), but during each brief interval of time reinforcement can occur independently of responding with probability \( p \). Clearly, the response is a better predictor of reinforcement when \( p = 0 \) than when \( p = .2 \), say. If operant behavior is sensitive to contingency, we might expect a higher response rate when \( p = 0 \) than when \( p = .2 \). In other words, adding “free” reinforcers should reduce the level of an operant response maintained by the production of response-dependent reinforcers (“welfare reduces work,” one might say), and this is usually the case (e.g., Hammond, 1980). Predictiveness can also be reduced by reducing \( q \), reinforcement probability: the lower the value of \( q \), the less well does a response predict reinforcement. In this case also, the less the predictiveness, the lower the response rate. In general\(^{98}\), the rate of the reinforced response is directly related to the difference \( q-\)

\(^{98}\) There are exceptions when \( q \) is very small (high ratio schedules) and regulatory processes come into play. Under these conditions, response rate may be inversely related to \( q \). See Chapter 7 et seq.
Reinforcement: Anomalous Properties

Autoshaping (Brown & Jenkins, 1968) occurs in hungry pigeons when they are occasionally (every 60 s or so) exposed to a briefly (7 s or so) lighted response key which is followed by food delivery. Pigeons never previously trained to peck a key nevertheless come to do so within a few trials – and will continue to peck (albeit at a lower rate) even if pecking turns off the light and prevents food on that trial (Williams & Williams, 1969). The experimental arrangement here is Pavlovian (the lighted key is the conditioned stimulus, the food the unconditioned stimulus), not operant (key pecking has no, or a negative, effect on food delivery). Since pecking is a “skeletal” response much used in operant conditioning, and since it was at that time widely assumed that only autonomic responses (like salivation in dogs) are susceptible to classical conditioning, autoshaping posed a problem.

Superstitious behavior posed a similar problem. Again, the procedure is Pavlovian, periodic presentation of a reinforcer (temporal conditioning, see Figure 10.8, a fixed-time schedule in operant terminology). Again, the effect is to produce skeletal behavior, despite the lack of any operant contingency – there is no causal relation between behavior and reinforcement (Skinner, 1948; Staddon & Simmelhag, 1971; Staddon, 1992).

Instinctive drift is the name given by Breland and Breland (1961) to their discovery that an animal trained to get food by making Response A will under some conditions after a while switch to (ineffective) Response B, even though it thereby misses food deliveries. For example, a raccoon may be trained by successive approximations (shaping) to put a wooden coin into the slot of a piggy bank (the Brelands trained animals for commercial advertising). But no sooner has it mastered this skill than it begins to mess up by “washing” the wooden tokens with its paws, as it might if the token were not wood but a piece of real food. Since the “washing” interferes with food delivery, and since the animal clearly “knows” how to get food, this instinctive drift, as the Brelands called it, poses real problems for standard learning theory. Since the behavior is also irrational by anyone’s definition, it also poses problems for almost any variety of behavioral economics (see Chapter 3).

**Figure 10.1** Structure of the model. Each activity has a strength defined by the output of an integrator (A–C for the three activities in the figure) whose inputs are reinforcement (R) and noise (ε: Equation 10.1 in the text). Outputs compete according to a winner-take-all (WTA) rule. This structure is very similar to the model for Stentor avoidance described in Chapter 4.
A REAL-TIME THEORY

Biologically valuable events – reinforcers – such as the occurrence of food or the availability of a mate, have two immediate effects on organisms: they arouse the animal so that it moves around more and engages in a variety of activities; and the arousal dissipates with lapse of time (Killeen, Hanson & Osborne, 1978). Most theories of adaptive behavior also accept that incompatible behaviors compete and the strongest usually wins (see the discussion of reflex interaction in Chapter 5). Finally, variability, the ability to generate a repertoire of different activities, is an essential ingredient of operant behavior, which is said to be freely “emitted” rather than “elicited” by a specific stimulus (at least in the initial stages of training).

The model I now discuss embodies these five properties – arousal, adaptation, strength, competition and variability – in a single linear discrete-time equation that applies to each activity in the organism’s current repertoire. The structure of the model is illustrated in Figure 10.1.

I define for each behavior a variable, \( V_i \), its strength. The response rule is winner-take-all competition: the activity with highest \( V \) value is the one that occurs. The equation describes the changes in \( V \) values from one time step to the next:

\[
V_i(t+1) = a_i V_i(t) + \varepsilon(1-a_i) + X(t) b_i V_i(t), \quad 0 < a_i < 1, \tag{10.1}
\]

where \( V_i(t) \) is the strength of the \( i \)th activity in discrete-time instant \( t \), \( a_i \) and \( b_i \) are parameters that depend on both the activity and the reinforcer, \( \varepsilon \) is a random variable sampled independently for each activity in each time instant and \( X \) is the value of reinforcement (typically 1 or 0).

In the absence of reinforcement \( (X(t) = 0) \), Equation 10.1 describes a low-pass filtered-noise process, since \( V_i(t+1) \) is just a weighted average of noise plus its previous value. Term \( a_i V_i(t) \) represents adaptation: because \( a_i < 1 \), this term reduces to zero with repeated iterations. \( a_i \) can also be thought of as a measure of short-term memory (STM): the larger \( a_i \), the more persistent the effects of any forced change in \( V_i \). The effect of \( a \) on the response of Equation 10.1 to an impulse input (i.e., a single fixed increment) is illustrated in Figure 10.2. The curves show that although the effect of the impulse is greater the smaller the value of \( a \), it persists longest at the highest \( a \) value.

Term \( b_i V_i(t) \) in Equation 10.1 represents

![Figure 10.2](image1.png)

**Figure 10.2** Response of Equation 10.1 to an impulse input at time step 10 for four values of parameter \( a \).

![Figure 10.3](image2.png)

**Figure 10.3** Selection property of reinforcement. In the simulation each time step is taken up with one of four identical (i.e., same parameter values) activities. Each curve shows the proportion of time taken up by the reinforced activity (reinforcement probability = 1) at a given pair of \( a \) and \( b \) values. Note that in the absence of selection, each activity takes up .25 of the total time. Each point is the average of \( 2 \times 10^5 \) iterations. Random variable \( \varepsilon \) in Equation 10.1 had a rectangular distribution over the interval 0-1 (reprinted from Staddon & Zhang, 1991, Figure 11.2).
the arousal effect of a hedonic event (reinforcement). Notice that reinforcement acts equally on all activities. The model handles both reinforcement and punishment: If \( b_i > 0 \), the effect is to increase \( V_i \) (a positive reinforcer); if \( b_i < 0 \), the effect is to reduce \( V_i \) (a punisher). Note also that the relation \( a_i + b_i < 1 \) must hold if \( V_i \) is not to rise without limit in repeated iterations of Equation 10.1 with \( X > 0 \) in each time step.

**Standard Properties**

**Selection.** If all activities are identical (i.e. \( a_i = k_a; b_i = k_b \) for all \( i \)), the model simulates all the standard properties of operant reinforcement. In the absence of reinforcement, because the two parameters are the same for both activities, each activity will occur equally often, on average. If reinforcement is delivered for each occurrence of Activity 1, then at that instant by the highest-wins competition rule \( V_1(t) > V_2(t) \). Because each activity is incremented in proportion to its value, then, neglecting the effect of noise, the increment to 1 must be greater than the increment to 2 so that \( V_1(t+1) > V_2(t+1) \). If the reinforcement occurs frequently enough that the increment in \( V_1 \) does not decay to zero by the next reinforcement \( V_1 \) will on average be incremented relative to \( V_2 \), so that Activity 1 will come to dominate. This conclusion holds whichever activity is reinforced; thus the process satisfies the reversibility condition.

The essential feature of the reinforcement mechanism is that immediate, response-contingent, reinforcement adds some increment to all activities, but the largest increment goes to the highest-\( V \) activity. Making each increment proportional to the \( V \) value is the simplest way to accomplish this, but other methods would probably give similar results. (Adding a constant increment to all does not work.)

**Figure 10.3** shows the steady-state effects of reinforcing every occurrence of one behavior, in a set of four, for a wide range of parameter pairs. The reinforced behavior is always facilitated, and the proportion of time taken up increases with increases in either parameter. The same thing happens for repertoire sizes of 2, 4 and 8, and thus presumably for any repertoire size.

**Delay.** **Figure 10.4** shows the effects of reinforcement delay. The higher the \( a \) value, the more likely the effective response will persist across the delay, and the less likely, therefore, that the delayed reinforcer, when it actually occurs, will strengthen some other behavior – so the shallower the gradient. But the delay-of-reinforcement gradient is relatively independent of the values of \( a \) and \( b \) so long as they sum to a constant.

Note that if two identical activities with high \( a \) values are both reinforced, but with a small difference in delay, then the activity that is more contiguous with reinforcement will eventually predominate. The system is potentially very sensitive to small contiguity differences. Animals are also, which poses a problem for the original explanation for “superstitious” behavior, as a failure to discriminate much larger contiguity differences. The suggestion is that the animal responds superstition by because it “thinks” its behavior is responsible for food delivery: “The bird [on a response-independent periodic food schedule] behaves as if there were a causal rela-
tion...[between its behavior and food delivery]” (Skinner, 1948, p. 171)  Skinner’s hypothesis leads to a paradox, given experimental demonstration that pigeons are capable of discriminating response-dependent from response-independent events with great precision (Killeen, 1978; Staddon, 1992).  I show in a moment that Equation 10.1 is consistent with both “superstition” and high sensitivity to contiguity, hence resolves this paradox.

Contingency.  Contingency is the fact that the strengthening effect of reinforcement depends on its correlation with the reinforced behavior, not just on contiguity.  Thus, if the target behavior is reinforced intermittently, or if it is reinforced every time but reinforcement also occurs at other times, the behavior will be strengthened less than if it is reinforced invariably or exclusively.  Figure 10.5 shows that the model has both these properties: Vertical comparison shows the effect of the probability of response-dependent reinforcement: the higher the probability, the larger the proportion of time taken up by the reinforced activity.  Horizontal comparison shows the effect of “free” (response-independent) reinforcers: the higher this probability, the lower the level of the explicitly reinforced act.

Thus, this system provides a simple qualitative account for all the standard steady-state properties of operant reinforcement.  The account is essentially parameter-free, since the pattern of results is the same for all parameter pairs, so long as all activities are identical.

Anomalous Effects.

Anomalies arise in the model when activities have different parameter values but in one case, “superstition,” apparently anomalous effects can arise even if all activities are the same.  I will discuss three anomalies: superstitious behavior, typologies, such as the distinction between emitted and elicited behavior, and instinctive drift.  Autoshaping involves stimulus-response associations and I defer discussion until the section on Pavlovian conditioning.

“Superstitious” behavior (Skinner, 1948; Staddon & Simmelhag, 1971; Timberlake & Lucas, 1985) is typically produced when an animal such as a hungry pigeon is given periodic response-independent food reinforcement.  If food delivery is frequent enough, the animal develops various kinds of vigorous stereotypies, despite the fact that food delivery is independent of its behavior.  The model is too simple to account for the temporal properties of this phenomenon – the fact that the activities occur at different times in the interfood interval, that there are two or three types of activity, etc. (see Staddon & Simmelhag, 1971, and Staddon, 1977, for details) – but the model can under restricted conditions produce apparently stereotyped behavior, even with a set of identical activities.

An example is shown as the four cumulative records on the left in Figure 10.6.  The curves show the levels of four identical activities, with high-a and low-b values, under continuous reinforcement.  The figure shows that over any given epoch on the order of 100-1000 iterations, the distribution of activities will be highly skewed, with one activity tending to predominate.  Examples are indicated by the vertical lines: over the epoch labeled “A”, for example, Activity 1 predominates; whereas during epoch B, Activity 2 is the dominant one.  Eventually, every activity will have its place in the sun.  But if the period of observation is limited, one activity will seem to have been preferentially strengthened.  The right panel in Fig. 5 shows the effect of
reducing reinforcement probability: the epochs over which behavior appears stereotyped are now much shorter. Under these conditions, there will be little temptation to cry “superstition,” even if the observation period is relatively brief.

The stereotypy illustrated in the left panel is the outcome of a positive-feedback process that resembles the original suggestion of Skinner (1948). Skinner’s explanation for all superstitious behavior is that when a (response-independent) reinforcer is delivered some behavior will be occurring and will be automatically strengthened; if the next reinforcer follows soon enough, the same behavior will still be occurring and will be further strengthened, and so on, until the behavior appears to dominate.

Two ingredients are lacking in Skinner’s account: an explicit statement about the temporal properties of behavior that are necessary for the process to work (the STM parameter, \( a \), in the present model); and explicit recognition of the importance of the time window over which observations are carried out. Dynamic analysis shows that (a) in this model the STM parameter must be very high for the process to work at all, even with reinforcements occurring at the maximum possible rate. (b) The time window is critical: if observations are continued for long enough, no activity will predominate. Thus, the effect of “free” reinforcers is simply to increase the autocorrelation of activities, rather than to single out one over others permanently (see Staddon & Horner, 1989, for a theoretical discussion of this issue in the context of stochastic learning models). And (c) even a modest reduction in reinforcement probability causes a sharp reduction in stereotypy. Since reinforcement is in fact intermittent in all published demonstrations of superstition, Skinner’s process, or at least in the present version of it, is unlikely to be responsible. On the other hand, Equation 10.1 is in fact a rather weak reinforcement mechanism, in the sense that it favors the reinforced activity only slightly if reinforcement is intermittent (see Figure 10.5). Given a more sensitive system (perhaps obtained by adding associative – stimulus-control – features to the simple scheme), “superstition,” in the sense of Figure 10.6, may be obtainable even under intermittent reinforcement.

Nevertheless, this analysis of superstition shows how the model accounts for the reliable finding of hysteresis in operant conditioning: the fact that once a response has been strengthened via response-contingent reinforcement, it may persist even when reinforcers are delivered independently of responding (Herrnstein, 1966; Rescorla & Skucey, 1969; see also Chapter 1). Given a high enough \( a \) value, even intermittent “free” reinforcers may be enough to maintain an already-strengthened activity, at least for a while, and for longer than if reinforcement were withdrawn entirely.

**Emitted and Elicited Behavior.** The distinction between emitted and elicited behavior parallels the procedural difference between operant and classical conditioning (Skinner, 1938). Elicited behavior (salivation is an example) is behavior elicited by a reinforcer but not modifiable by operant conditioning. Emitted behavior (lever pressing is an example) is not usually elicited by the reinforcer, but is modifiable by operant conditioning (the rat learns to press the lever to get food).
The model suggests why this distinction has some empirical validity. Recall that the two parameters in Equation 10.1, \( a \) and \( b \), cannot sum to more than one if the system is to be stable under all conditions. Moreover, \( a \) determines how effective operant reinforcement will be: small \( a \) means that the effects of reinforcement will not persist, no matter how large they might be initially (cf. Figure 10.2). Conversely, a behavior that gets a big boost from reinforcement, because of a high \( b \) value, cannot have a high \( a \) value, hence will not show good operant conditioning. Consequently, activities that are elicited by reinforcement are unlikely to be good candidates for operant strengthening: elicited (Pavlovian) and emitted (operant) behaviors are likely to form separate categories – although the categories are also likely to overlap, because the distinction is at bottom quantitative, not qualitative. But it is a powerful implication of the simple model that elicited and emitted behavior should have different properties, because of the complementary relation between the arousal \( b \) and STM \( a \) parameters that is forced by stability considerations.

**Instinctive Drift** is perhaps the most striking exception to a contiguity account of reinforcement. One explanation for the behavior involves two steps (Staddon & Simmelhag, 1971). The first step involves purely operant reinforcement – of placing the token in the slot, in the raccoon example, say. Once the animal has learned this response, the token becomes a reliable predictor of, and is reliably paired with, the delivery of food. This sets the stage for the second step, which is Pavlovian. The token becomes a conditioned stimulus for food-related behaviors. Since the token resembles food in its physical properties, the raccoon begins the “washing” movements that it characteristically shows to bits of food. Since it cannot both “wash” and insert the token in the slot, this behavior prevents food delivery, which leads to extinction of the “washing” and eventual recovery of the effective, operant response. The cycle then repeats.

This is a perfectly plausible associative account, but Equation 10.1 also suggests a completely non-associative explanation for instinctive drift. The non-associative account derives from the possibility that different activities may have different \( a \) and \( b \) values. Suppose that “instinctive” behaviors are ones that have high \( a \) values. Consider then the situation where the animal’s repertoire consists of a number of activities with moderate \( a \) and \( b \) values but a single activity with a high \( a \) value. Operant reinforcement of one of the moderate activities will cause it to predominate for a while. But, because increments in the \( V \) value of the “instinctive” activity cumulate more effectively than the (larger) increments to the other activities, it may predominate eventually, even though it is never contiguous with reinforcement.

These effects are illustrated in the cumulative records in the top panel of Figure 10.7. Each set of four records shows four activities, three with moderate \( a \) and \( b \) values, one with a high \( a \) value and low \( b \) value. The Left records show the operant (unreinforced) levels of the four activities; notice the low frequency of the high-\( a \) activity. (Although the form of Equation
10.1 ensures that all activities will have the same average \( V \)-value, the variance of the \( V \)-values will be inversely related to \( a \). Thus, a single high-\( a \) activity will occur less often on average than an activity with lower \( a \)-value when there is more than one low-\( a \) activity.) The Center and Right panels show the effect of reinforcing one of the low-\( a \) activities with a probability of .25 or 1.00. The increasing reinforcement probability has two effects: it causes the level of the reinforced activity, “\( R \),” to increase above the level of the other two low-\( a \) activities; but it causes a disproportionate increase in level of the high-\( a \) activity, which becomes predominant when \( p(R) = 1 \) (Right). The two bottom panels in Fig. 10.7 show a magnified picture of initial acquisition in the \( p(R) = 1 \) condition, illustrating the transition from dominance by the reinforced, low-\( b \), activity to predominance of the “instinctive,” high-\( a \), activity. The message is that if a low-\( a \) activity is reinforced with high enough probability, it may predominate initially, but may eventually be supplanted a high-\( a \) activity, despite its lower “arousability” \( (b \) value). If “washing” in the raccoon example is a high-\( a \) activity, but putting the token in the slot has a lower \( a \)-value, then washing will eventually win, even if putting the token in the slot is the reinforced activity.

**Unsignaled Shock-Avoidance**

Avoidance behavior in the absence of a “safety” signal has always posed a problem for the contiguity view of reinforcement because there is nothing for the avoidance response to be contiguous with. An explicit reinforcing event is completely lacking in what is termed a shock-rate-reduction schedule, for example (Herrnstein & Hineline, 1966). In such a schedule brief shocks are delivered with probability \( p \) during each sampling interval, \( \Delta t \) (\( \Delta t \) is usually on the order of a second or two). If the animal makes the designated avoidance response, \( p \) is reduced to a value \( p' < p \) until the next shock occurs, when it returns to \( p \). Rats learn (albeit with some difficulty) to make the avoidance response on such procedures, despite the fact that no tangible event occurs contiguous with the avoidance response. The model can accommodate this behavior if it is assumed that electric shock and other aversive events reduce, rather than increase, \( V \) values (i.e., \( b < 0 \)). Under these conditions, electric shocks reduce all \( V \) values, but the \( V \) value of the avoidance response will be reduced less than others’ because on average it is followed by shock after a longer delay. Hence, it will be favored in the competition and, in an ensemble of identical competing activities, will come to predominate.

**PAVLOVIAN ASSIGNMENT OF CREDIT**

On the Indonesian island of Bali, the people grow rice in water-filled terraces. The pattern of cultivation is ancient — the ‘Subak’ irrigation system itself is a miracle of evolved coordination — and nothing is wasted. Each rice paddy is tended by several villages, and after the rice is harvested, flocks of domestic ducks are brought in by ‘duck women’ to feed on the gleanings that remain. As dusk falls, each woman leads her ducks back to their village so they can spend the night safe from predators. How, one might ask, does each duck woman identify her own ducks? How does she separate hers from others’? How does she herd them?
The answer is: they herd themselves. The scheme works like this. Before a duck is allowed to forage freely it first learns the signal for food, usually a flag waved to announce feeding time each day. The birds learn the association quickly. Each village has a flag of a different color or pattern. When each duck woman needs her flock, she waves her flag. Only her birds recognize the flag and only her birds come when she waves it. In this way, the ducks of each village sort themselves out every evening. This is an ancient example of what is now known in the trade as Pavlovian or classical conditioning.99

Much less picturesquely, Pavlov showed early in the twentieth century that animals can learn to anticipate food if it, the unconditioned stimulus (US), is reliably preceded by a signal, the (to-be) conditioned stimulus (CS). The dog will salivate (conditioned response: CR) to the tone (CS) if the tone has been reliably followed by food (US). Three standard arrangements are illustrated in Figure 10.8. The study of classical conditioning has proliferated mightily in the nearly 100 years since Pavlov’s original work. The effects of CS duration, of CS-US delay, of CS probability, of number and history of CSs serially and in compound, of CS type and US type, of different CS-US combinations, and dozens of other experimental arrangements have been extensively explored in dogs, cats, rats, rabbits, sea slugs, blow flies and many other species. The properties of conditioning are remarkably similar in these different organisms, suggesting that conditioning reflects universal regularities in the causal structure of the world, rather than neural structures that are homologous from flies through cephalopods to humans100.

Research on classical conditioning has shown (among other things) the following:

1. Delay. CS-US delay matters: the longer the delay between CS onset and the onset of the US, the smaller the conditioning (CR strength, measured in the various ways I have already discussed for reflex strength). See also the ratio rule, below.

2. Contingency. Probability – US|CS and US|~CS – matters. It is important not just that the CS just-precede the US, but the pairing must be reliable: not too many CSs that are not followed by the US; and the US must not occur too often on its own. There must be a CS-US contingency: p(US|CS) > p(US|~CS).

---

99 This is a lovely story, but of course no one has done experiments with Balinese women and their ducks to isolate the critical variables. It is possible, therefore, that at least two other well-known processes may be involved. One is imprinting, the fact that ducks and other precocial birds will learn to follow for the rest of their lives something to which they are exposed during the first couple of days after birth. The “something” may be mother duck, but it may also be a flashing light or a colored object – like a flag. The other process is following. Ducks, like sheep, goats and cattle, are social creatures and prefer to stay together. Hence a whole flock may follow even if only a few animals become conditioned and go after the flag. As is often the case in adaptive behavior, many processes may conspire to produce a given result.

100 Although we cannot be quite as confident of this today as we might have been several years ago, given the extraordinary genetic continuities between organisms as distantly related as fruit flies and humans that have been revealed recently. See, for example, Bull & Wichman (1998) and numerous other recent articles in Science and Nature.
3. **Conditioned Inhibition.** If a stimulus reliably signals absence of the US, it acquires inhibitory properties: the CR is suppressed in its presence, and it may reduce the level of CR elicited by other CSs if it is presented with them.

4. **Latent inhibition.** If a potential CS is presented repeatedly on its own, before conditioning trials begin, it becomes harder to condition than a totally novel CS. This is termed latent inhibition.

5. **Sensory preconditioning.** If two stimuli, A and B, are repeatedly presented in succession – AB, AB, AB, etc. – then in the second phase of the experiment, B is repeatedly paired with a US in the usual Pavlovian way, then in a test with stimulus A alone, there may be some evidence of conditioning to A, despite the fact that it has never been either paired with, or even in the same context with, the US. This is termed sensory preconditioning.

6. **Overshadowing.** If there are two coextensive CSs – simultaneous tone and light, for example – with the same temporal relation to the US, one may nevertheless condition more than the other. The more salient stimulus *overshadows* the less salient.

7. **Blocking.** But, given two equally salient stimuli, A and B, if stimulus A gets several pairings with the US before stimulus B is introduced, making an AB compound stimulus, B will get less, or even no, conditioning. In a subsequent test A will evoke the CR, but B may not, despite several AB-US pairings. A *blocks* B.

8. **Trace and temporal conditioning.** Conditioning can occur even if the CS does not overlap the US (trace and temporal conditioning, see Figure 10.8). In these cases, the CR occurs after the CS, in anticipation of the US. Temporal conditioning, in which the US is also the CS, is easier to obtain than trace conditioning, in which the CS is a neutral stimulus. A US makes a better *time marker* than a neutral stimulus.

9. **Serial conditioning: Occasion setting.** Consider two CSs in series, as in the top of Figure 10.9. A CR will develop to both stimuli, but their properties are very different. For example, suppose that after conditioning, CS\(_1\) in isolation (as in delay conditioning, see Figure 10.1) elicits CR\(_1\) and CS\(_2\) in isolation elicits CR\(_2\). Nevertheless, in the serial arrangement, CS\(_1\) may now elicit quite a different response, CR\(_x\). Moreover, if trials like the top sequence are alternated with trials like the bottom – so that CS\(_1\) is a signal that CS\(_2\) will be followed by the US, CS\(_1\) takes on very different properties than when it signals the US directly. It “sets the occasion” for CS\(_2\) to act as a CS and rather than being a simple excitor itself. A vigorous experimental program has been provoked by the differences between occasion-setters and standard conditioned stimuli (cf. Holland, 1992; Swartzentruber, 1995).

10. **Autoshaping and the ratio rule.** Pavlov, interested in digestive processes, not psychology, studied salivation as the conditioned response. It was at one time thought that only autonomic responses of this sort are subject to classical conditioning. But Brown and Jenkins (1968), thinking they had simply devised an efficient way to train hungry pigeons to peck a key, showed that pecking and, perhaps, other skeletal, “voluntary” responses could be classically conditioned. It turns out that pairing a CS with a US can produce and maintain a wide variety of “operant” behaviors (Staddon & Simmelhag, 1971; Staddon, 1977). In systematic studies of autoshaping, Gibbon, Balsam and their associates showed that the most important determiner of conditioning is the ratio of ITI to CS duration: the larger the ratio – the more predictive the CS – the better the conditioning (Gibbon & Balsam, 1981).

Some of these effects are much more complex than others. But all are aspects of *predictiveness*, which is fundamental to classical conditioning. Blocking, for example, shows that if a US is already well predicted by one stimulus, adding a second yields little or no additional conditioning. The ratio rule shows that a CS that is brief, in relation to the time between US presenta-
tions, will get more conditioning than a long CS. The fact that the CR tends to occur after the CS, and close to the time-of-arrival of the US, in temporal and trace conditioning, shows that post-event time may itself act like a CS, and times close to the US are more predictive, hence attract more CR, than times remote from the US. The standard measure of contingency, \( p(US|CS) > p(US|\sim CS) \), is also the standard measure of information transmission in signal-detection analysis (e.g., Green & Swets, 1989).

The Dynamics of Classical Conditioning

Classical conditioning has for some time been taken as the prototype for simple learning, which is why it has been studied in numerous lowly species, from sea slugs to blowflies. But in fact, when one attempts to put together a simple dynamic account that can accommodate if not all at least most of the ten phenomena just described, it is hard to avoid the impression that the process is in fact very complicated. A reasonable model, one feels, should accommodate the complex relation between classical and operant conditioning that was so powerfully highlighted by the discovery of autoshaping. It should also say something about the transformation of the UR into the CR and the fact that these two responses are often very different in form. It should deal with occasion setting, which shows very different stimulus and response properties from simple delay conditioning. It should accommodate the acute dependence of all conditioning phenomena on temporal variables – CS-US delay, the durations of stimuli in serial conditioning, the duration of the intertrial interval and several others. It should account for the existence of conditioning in the absence of any US (sensory preconditioning). Yet in fact existing conditioning models deal with only one or two of these effects at a time.

Successful assignment of credit always depends on some kind of competition. The details vary, but the common theme for all the mechanisms that have been proposed for classical conditioning is that each of the different potential CSs (stimuli) is assigned a number that reflects (for example) its proximity in time to the US. Then some kind of learning process changes the number in a non-linear way that favors the event closest to the US. The end result is that the stimulus that best predicts the US (in the limited sense of temporal proximity and frequency of co-occurrence) gets the highest number and thus the most conditioning.

The earliest scheme of this sort is the Rescorla-Wagner model (Rescorla & Wagner, 1972). The R-W model is not real-time (real-time version was proposed by Ludvig, Sutton & Kehoe, 2012); it operates trial-by-trial and takes no account of temporal relations. But it can detect contingency and does show the phenomenon of blocking. Competition in the R-W model takes the form of a fixed maximum associative strength, set by the prevailing conditions of reinforcement, for which all stimuli must compete. The idea is that each conditioning trial increases the associative strength of both relevant and irrelevant stimuli (usually termed “CS” and “con-...
text”) and each extinction trial (the intertrial interval, for example) reduces all associative strengths. Because relevant stimuli are present only on conditioning trials, they use up all the available associative strength and attract all the conditioning. Blocking is explained because pre-training with CS A allows it to attract all the associative strength before CS B gets a chance to compete. When B is presented, therefore, there is no remaining associative strength, and B fails to condition.

The R-W model is deficient in certain respects in addition to the fact that it is only a trial-level model. Most notably, it cannot account for latent inhibition. A simple model that can do so has been proposed by Pearce and Hall (1980) and much more elaborate neural-net models that explain not just contingency, blocking, etc. but also many of the more “cognitive” properties of classical conditioning in animals (and perhaps some aspects of their physiological substrate), have been proposed by Grossberg, Kehoe, Miller, Schmajuk and others (e.g., Grossberg, 1987; Kehoe, 1998; Miller & Matzel, 1988; Schmajuk, 1997).

Classical conditioning, unlike the simple operant assignment-of-credit scheme discussed in the first part of the chapter, is associative; that is, it implies some kind of connection between a specific stimulus and a response. I have not so far dealt with associative processes, because of their evident complexity. But the next chapter is a beginning attempt to deal with one aspect of association, stimulus control, in the context of spatial navigation.
Chapter 11

STIMULUS CONTROL

A recently discovered addendum to *The New Atlantis*, the great work of scientific philosophy by philosopher, courtier and political schemer Sir Francis Bacon (1561-1626), contains a fragment of dialogue between two philosophers, Polaris and Arcturus, about the nature of the world:

*Polaris:* But, Sir, I do believe the elements of the Ancients, Earth, Air and two others, can not take in the motions, colors and humors of the celestial orbs. Mercury, in his lively motion, and even torpid Jupiter, must need *vis viva* for a full account. I argue for a vital, quicken’d spirit as organizer of the World.

*Arcturus:* It cannot be, sage Polaris. Four elements are given, divine, and must do for all. I know that some have more proposed, five or even (or is it odd) seven, number beloved of saints and necromancers. But then where can all end? What number is right – if not four? *Vis viva* is a vap’rous, insubstantial notion which can be no true philosophy.

*P:* Esteemed and able Arcturus, surely no elements, four or four hundred, can together vitality infuse. None can in combination more than either gain in movement, life and spirit. Fair Mercury, fast and bright, scudding o’er the scorching sun, *must* feel his might and ever seek escape. Air and water, earth and fire be not enough.

*A:* Polaris, thou ignorant wizard wannabe, Of four elements, all together, more than each may easily come out. What of this *vis*, and where *viva*? What doth it do? What can it not do? Where is the principle of Occam, long forgot by some, but sage and sovereign in all philosophy? It is vain to do with more what can with less be done…

and here the fragment ends. This is fiction, of course; Bacon never wrote these words. But, as the flexi-truth ideologues say, ‘It *could* have happened…’ Indeed, it *is* happening in a debate rumbling as I write between devotees of two opposed approaches to the study of learning: association and computation\(^{102}\). The associationists argue that all learning can be accounted for by the formation of stimulus-response links. The computationists, energized by the model of the digital computer (and the successes of the ‘cognitive revolution’), argue that if not all, at least all the most interesting, examples of learning must be explained as symbol manipulation, a computational process.

The problem with this debate is that like the *vis viva* of Polaris and the ‘elements’ of Arcturus, neither association nor computation is well defined, so they do not exhaust the universe of theoretical alternatives. We may suspect that there are useful models that are neither associationist nor computational. To frame the debate as a dichotomous choice between ill-defined alternatives, too much constrains the theoretical imagination.

Spatial navigation, the main topic of this chapter and the next, has traditionally been considered a cognitive — hence computational — topic, because of the apparent subtlety of animals’ spatial behavior. The approach I will discuss is closer to the associative than the computational pole. But really it fits neither category very well. The general point is that the issue cannot be decided in the abstract. It is probably premature to reach for conclusions about the metatheory of learning before we have some viable theory. Better first to come up with valid models for specific sets of data. Here is a beginning approach to spatial navigation.

---

GENERALIZATION and SIMILARITY

Bacterial orientation (Chapter 4) is a kind of operant behavior, but it is limited in a crucial way. Bacteria solve the operant assignment-of-credit problem (What must I do?), but seem not to deal with Pavlovian assignment of credit (When must I do it?), a more complex task. “Higher” animals can learn not only what responses to make but when — in the presence of what stimuli, at what times — to make them. The properties of the system they use for dealing with stimuli have been studied under the headings of stimulus control, stimulus generalization and interval timing. This chapter and the next deals with one aspect of stimulus control/generalization; Chapters 13 and 14 are devoted to interval timing.

Stimulus Generalization

After behaving in a certain way in the presence of one stimulus, animals tend to behave in the same way in the presence of other, similar stimuli, they generalize from one stimulus to another. The tendency to generalize can be measured in various ways. For example, a hungry pigeon is first trained to peck a response key for food reinforcement, delivered on a variable-interval schedule of intermediate value (e.g., VI 60 s). The response key is illuminated with the stimulus of interest (termed $S^+$ or $S_D^+$). $S^+$ is usually chosen to be physically simple — light of a single wavelength, a single vertical line, a tone — but physically more complex stimuli, such as pictures of scenes or animals can also be used. After the pigeon has learned to peck $S^+$ for food, variants on the training stimulus — stimuli that differ from $S^+$ on a physical dimension such as intensity or wavelength — are presented for relatively brief periods. For example, if five test variants are tried, each may be presented for one minute in randomized blocks of five, for a total of perhaps 60 presentations. Each variant is presented repeatedly, so that unsystematic variations in the tendency to respond are shared equally by all. No reinforcement occurs during a generalization test, so as to avoid reconditioning. Extinction is mitigated by the VI training: on VI, long periods without food are common. Hence, VI-trained 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 $S^+$.

Generalization gradients measured in this way tend to be sharply peaked (if the number of test stimuli is few and they are spaced far apart), but bell-shaped if more, closely-spaced stimuli are used. Gradients seem to be sharper when the training schedule is rich (frequent reinforcement) than when it is poor (low reinforcement rate). In every case, response rate is maximal at the training stimulus value and declines smoothly at values above and below. A great many physical stimulus properties — wavelength of light, line tilt, roughness, spatial frequency and others — have been tested like this. The universal result, with pigeons, rats, monkeys, goldfish and people, is that responding is maximal at (or near) $S^+$, and falls off systematically as the physical stimulus difference between $S^+$ and the test stimulus increases.

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. But in other cases, the result is puzzling. Color perception, for example, depends upon central integration of information from three or four types of photoreceptor, each with a different wavelength-absorption peak. Wavelength, as a continuum, is not simply represented in such a system. Yet wavelength generalization gradients are among the most orderly, and show no sudden slope changes, even at color boundaries. What is true for wavelength, deeper examination shows to be true for all stimulus dimensions: generalization gradients measure neither physical differences nor the properties of peripheral sensory transduc-
Similarity. Generalization gradients are measures of similarity, a psychological property that relates not so much to perception, as to action. The evolutionary point of generalization is not to group stimuli according to their physical resemblance, but to treat as similar stimuli that signal similar consequences. The fact that similar stimuli are often physically similar is either a happy coincidence, or an interesting property of the physical environment, as you wish. Similarity in this sense is assessed by the organism in two ways. By training: stimuli associated with the same reinforcement schedule are responded to similarly. By nature: stimuli that in the history of the species have been reliably associated with similar consequences may lead to the evolution of innate equivalencies. The most obvious example is our perception of an object. Different views of the same object share few physical properties (although they are geometrical transformations of one another). Yet even infants treat different views of the same thing as similar. The perceptual constancies, in which objects appear the same even though their sensory properties have changed, are another example. Learning plays only a minor role in the development of these processes, which are largely built-in.

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 dissimilarity 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.
There are two ways to create a similarity space: the simpler method takes physically defined objects and transforms their physical dimensions so that the transformed values reflect similarity, as shown by the fact that after the transformation, generalization gradients have the same form, irrespective of where they lie on the stimulus dimension. It is not always possible to reduce gradients to an invariant form through a simple transformation, however. In this case, the second method must be used. A similarity space is derived by multidimensional scaling (MDS) techniques from empirical similarity judgements. The dimensions of such a space need not be simply related to physical dimensions. These two approaches are most easily understood through examples.

The first type of similarity space is shown in Figure 11.1. The top panel shows the results of a hypothetical experiment in which generalization gradients were successively obtained with three different S+s (arrows). The figure shows that the gradient spread is directly related to the value of S+: the same stimulus difference is associated with a large change in response strength in the test with the smallest S+, a much smaller change with the largest. This difference implies that the psychological effect of the same physical difference is substantial at the low end of the physical scale, but less impressive at the high end. This dependence of differential sensitivity on base value is characteristic of many physical dimensions, notably sound and light intensity, as well as to judgements of time: a second is a lot when judging intervals on the order of three or four seconds, it is not so much when judging intervals on the order of minutes. For most intensive dimensions, the ratio between a measure of response variation (such as the standard deviation of the generalization gradient) is proportional to its mean over a considerable range: this is one implication of the Weber-Fechner law, that the just-noticeable difference is proportional to the base value. If a 1-lb increment is noticeable 50% of the time when added to a 5-lb weight, then 2 lb will be just-noticeable when added to a 10-lb weight.

The lower panel of Figure 11.1 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. Some measures of interval timing show exactly this property: skewed generalization gradients if the data are plotted against real time, symmetrical gradients if log time is used (see Chapters 13 & 14). Notice that the log transformation here says nothing about the actual form of the generalization gradient, which might be determined by several factors. It implies only that gradients obtained under comparable conditions with different S+s will be rendered identical in form by log-transforming the stimulus axis.

Figure 11.1 illustrates how the unidimensional world of physical intensity can be transformed into its psychological representation. The hypothesis of an invariant psychological “similarity space” can also be applied to more complex stimulus properties, such as color, where the physical and psychological dimensions are not simply related by a scale transform. The method is straightforward in principle. For example, given three stimuli A, B, and C with judged difference (1/similarity) relations AB = 1, BC = 1, AC = 2, the psychological space has just a single dimension, since the differences among the three stimuli are the same as those among three equidistant points along a line. But if the judgements had been, say, AB = 1, BC = 1, AC = 1.5, two dimensions would have been required, with the 3 points arranged in a triangle.

Color is an example like this. Numerous experiments have obtained judgments from people of the subjective similarity of color samples. In an old 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. But 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 — a
form of *color circle*. Thus, although the stimuli vary in only one physical dimension, wavelength, the psychological 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: hues, around the rim of the circle, and saturation, varying radially, from white in the center to highly saturated at the rim.

The color space is 2-dimensional, but the physical dimensions of stimuli are not simply represented in it: “north-south” does not correspond to one physical dimension and “east-west” to another, for example. Its essential property is that it accurately represents the invariances in a set of behavioral data.

In both these examples the “objects” dealt with are 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 similarity-space approach can work with essentially any object, even (perhaps especially) “natural” objects such as color pictures of actual scenes.

**Cognitive Maps.** Practical difficulties have meant that with one exception, rather little work on similarity spaces has been done with animals. The exception is work on orientation, how animals find their way about the physical environment. 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. A cognitive map is of course just another similarity space; in this case, one that closely resembles its physical counterpart. The function of particular stimuli is not so much to directly elicit approach or avoidance as to tell the animal where it is in relation to its map, that is, to function as landmarks.

In the laboratory, experiments with experienced animals in mazes show that visual cues outside the maze do in fact act as landmarks rather than as stimuli to be approached or avoided in push-pull fashion. For example, in a famous series of experiments initiated by Olton (e.g., Olton & Samuelson, 1976), rats were 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 visited.

In another ingenious experiment (Suzuki, Augerinos, & Black, 1980), a radial maze was covered with a black cloth dome on to 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.

Given a cognitive map, some active process of map-reading is needed to actually guide action. The dynamics of stimulus generalization suggest a simple process that can do much of the job.

**Dynamics of Stimulus Generalization**

Roger Shepard (1987) has proposed a “universal law” for stimulus generalization in psychological space. His idea is that physical dimension are mapped into psychological dimensions in a way that generates a “universal” similarity space. This is an ambitious claim for which evi-
dence is sparse, outside of a handful of physical dimensions. And there seems to be no particular evolutionary reason why such a universal space should exist at all. Nevertheless, Shepard’s arguments do suggest a useful approach to the problem of generalization dynamics.

Shepard offers both empirical evidence and functional (optimality) arguments in support of an exponential generalization function. However, Gaussian functions are found experimentally at least as often as exponential ones. Shepard (1988a&b) and Ennis (1988a&b) have both proposed additions to Shepard’s initial proposal aimed at reconciling Gaussian and exponential data.

Shepard’s original generalization theory is based on the concept of diffusion: “According to that model, on the removal of an external stimulus, the memory trace of the stimulus spontaneously undergoes a continuing process not only of simple weakening or decay but also of outward diffusion in psychological space (1958, p. 415).” Diffusion, of course, yields the Gaussian response-probability distribution, which will be increasingly broad as the time of testing is delayed after stimulus presentation. In the Shepard model the exponential function is derived by integration across previous Gaussian traces: “...under regularly repeated presentations of a stimulus, the integrated strength of the accumulated traces left by all preceding presentations approximates an exponential decay function of distance from that stimulus in psychological space (1958, p. 415).”

It is in fact possible to implement the diffusion idea in an even simpler way than Shepard’s proposal. Moreover, the method is capable of yielding both Gaussian and exponential gradients and, as I show in the next chapter, it also provides a process adequate to guide movement using a cognitive map.

The theory I describe (Staddon & Reid, 1990), like Shepard’s, is designed for an “...idealized generalization experiment in which an individual is given a single reinforced trial with a novel stimulus...and then is tested with another stimulus from that domain (Shepard, 1986, p. 60).” Suppose that the internal representation of a single stimulus dimension is as a line of units, each with four connections (Figure 11.2): two bidirectional connections to its two immediate neighbors and an input from a perceptual mechanism that allocates a limited region (e.g., wavelength band) of a single sensory dimension to each unit: when a given region is present in the stimulus, the corresponding unit receives a positive input. Each unit has as its final output an “activation strength”, $x_i$, that is its contribution to the measured generalization gradient.

Suppose that the essential property of this very simple locally connected net is that at each time step the activation of each unit moves towards the average of its neighbors’. Formally, the change in the strength of the $i$th unit in a series will be given by

$$
\Delta x_i = a[(x_{i-1} - x_i) + (x_{i+1} - x_i)], \quad 0 < a < .5,
$$

(11.1)

where $\Delta x_i$ is the change in $x_i$ from one discrete-time iteration to the next (i.e., $x_i(t+1) - x_i(t)$) and the term in parenthesis is the net strength difference between unit $i$ and its neighbors, $i-1$ and $i+1$. Equation 11.1 is a discrete-time version of Fick’s first diffusion equation (see, for example, Berg, 1983), and $a$ is a diffusion-rate parameter. During any iteration a unit may also receive stimulus input, so that the net change in activation of a stimulated node will be $\Delta x_i + S_i(t)$. Thus, the whole system can be represented as follows:

$$
x_i(t+1) = x_i(t)[1-2a] + a[x_{i-1}(t) + x_{i+1}(t)] + S_i(t),
$$

(11.2)
where \( S_i(t) \) is the stimulus input to the \( i \)th unit during iteration \( t \). (An alternative which turns out to be better for most applications is to assume that stimulus input activates a particular node, whose value is then determined by the presence or absence of reinforcement. The values of inactive nodes are determined by the diffusion process. I get to this modification in a moment.)

Figure 11.3 illustrates the behavior of Equation 11.2. The figure shows the activation gradient at various points following a stimulus at Unit 11 for a single time step at time zero. The spreading gradients show how the activation of the nodes changes across five iterations after a brief stimulus is withdrawn. After one time step the gradient is exponential, but then it collapses into a Gaussian form over the next four time steps. Thus, a diffusion process alone shows both the negatively accelerated and Gaussian gradient forms, depending on the time delay between target offset and the onset of the test stimulus.

Although the gradient is approximately exponential when a stimulus is present for a moderate number of iterations, the true steady-state form of the function is not exponential. It approaches a straight line under the following two conditions. If the borders (here, the end units) of the net are absorbing, that is, their activation value is clamped at zero; or if we look only at a finite region of the activation profile.

The present scheme has both advantages and disadvantages relative to Shepard’s (1958) model. The main disadvantage is that it involves short-term memory (STM) only, hence cannot explain retrieval effects that depend on stable stimulus representations (i.e., long-term memory: LTM). The advantages are that it involves just a single well-understood process; it relates in a natural way to connectionist theory; and it can be implemented by a very simple network. With slight changes, the rate of spread and the rate of decline in peak value of the gradient, which are both determined by the same parameter in simple diffusion, can be dissociated without changing the other useful properties of the process. This scheme can also easily be extended to more than one dimension by postulating nearest-neighbor connections in nets of two, three or more dimensions. It is easy to derive the effect of complex (multi-unit) stimuli on the generalization profile.

MDS analysis of experimental data by Shepard and others implies the existence of an invariant form of internal representation in the sense that the shape of the internal gradient does not depend on the physical value of the peak stimulus. The diffusion process has this property in the sense that apart from boundary effects the shape of the gradient, with respect to the net of units, does not depend on which unit is at the peak. Different forms of equal-generalization contours for multidimensional stimuli can perhaps be accommodated by different diffusion rates in different directions within the net. Although this scheme is useful for explaining generalization effects, its most powerful application is to spatial orientation.
SPATIAL ORIENTATION

The cognitive behaviorist Edward Tolman (1886-1959) spent much of his career devising clever experiments to show that stimulus-response (i.e., associationist) accounts of rat behavior cannot be correct. Some of his most striking demonstrations involve spatial learning. Figure 11.4 shows one such example, a maze apparatus used in a famous experiment by Tolman and Honzik (1930). The maze has three paths from the Start box to the Goal box. The paths differ in length: Path 1 (heavy vertical line) shorter than Path 2 (intermediate line) shorter than Path 3 (light line). In preliminary training, the rats were allowed to become familiar with all three paths to the Goal. They also had experience with a block at point A, which permits access to the Goal only via Paths 2 and 3. In the test condition, the block was moved to point B – so that only Path 3 is open. The question is: Will the rats choose Path 3 as soon as they encounter the block at B, or will they choose Path 2, which is normally preferred to Path 3 – indicating they do not know that Paths 1 and 2 share a common, blocked, segment? Most of Tolman and Honzik’s rats behaved intelligently and went straight to Path 3 after encountering the block at B. Tolman took this as evidence that the rats knew something about the topography of the maze. They were not just operating on a fixed hierarchy of preferences (“Path 1 better than Path 2 better than Path 3”) nor were they responding reflexively to local cues. Tolman considered this behavior to be an example of “insight,” although he did not specify exactly what that means. He did say that some kind of cognitive map must be involved, and he also argued that these results are incompatible with a blind S-R process. But he offered no satisfactory alternative, prompting Guthrie’s famous objection that Tolman’s theory left the rat “buried in thought” at each choice point (1935, p. 143; see also Dawkins, 1982). The problem was that Tolman assumed a map but provided no way for the animals to use the map to decide on where to go – no route-finder.

There have been several responses to Guthrie’s objection. For example, Deutsch (1960) described a verbal model for learning embodied in 14 propositions. “Insight” learning in the Tolman and Honzik apparatus is accounted for by assuming three “links” corresponding to the three alternative paths. When Path 1 is blocked, it is the link corresponding to Path 3 rather than that corresponding to Path 2, that is activated, because of common elements between Path 2 and the blocked Path 1. Deutsch did not discuss barriers, open-field navigation or cognitive maps. More recently, Hampson (1990) has sketched an ambitious connectionist program intended to model Tolman’s maze results and many others in the animal learning literature. I now describe a much simpler route-finding model (Reid & Staddon, 1998), based on the dynamics of stimulus generalization. I show that the Tolman and Honzik problem, and several other “standard” problems of spatial orientation, can be solved by dynamic stimulus generalization in an internal map.
The process involves no overview of the problem as a whole, hence nothing corresponding to “insight” — it requires a cognitive map, but makes do with a non-cognitive route finder. The analysis shows that the processing requirements for behavior that shows spatial “insight” are simpler than previously thought.

**NAVIGATION USING DYNAMIC GENERALIZATION**

Since Pavlov (1927), some kind of spreading activation has been a common element in theories of adaptive behavior. The model I now describe is a diffusion model but, unlike Pavlov’s “irradiation”, it is intended to explain behavior rather than brain physiology. I first describe one- and two-dimensional versions of the model. Both are modest extensions of diffusion-generalization.

**One-Dimensional Model**

The one-dimensional diffusion model assumes that a single stimulus dimension can be represented as a line of units (cf. Figure 11.2). Each unit has connections to its immediate neighbors. Associated with each unit is an activation strength, $V_i$, termed reinforcement expectation, for reasons that will become apparent. In the absence of external stimulation, expectation diffuses from one unit to the next according to an elementary diffusion process; that is, if $V_i > V_{i+1}$, then $V_i$ will decrease in the next time step and $V_{i+1}$ will increase (Equation 11.1).

In the absence of external input, this process will eventually lead to flat (or zero, if boundaries are absorbing) expectation across the whole network. This does not happen in the model, because expectation is injected into the network through active units. In the case of spatial generalization, one and only one unit will be active at any instant – the animal can only be in one place at a time. The value (expectation) of the active unit depends on whether or not reinforcement (e.g., food, for a hungry animal) occurs during that time step. If it does, the value of the unit is set to a positive value (e.g., 1, for unit reinforcement); if there is no reinforcement, value is set to zero.

Formally, in discrete-time notation, $V_i(t)$ is the expectation associated with unit $i$ at time step $t$. At each time step the expectation of each inactive unit moves towards the average of its neighbors’. The change in the value of the $i$th unit in a series at each time step is given by

$$ \Delta V_i = \alpha [(V_{i+1} - V_i) + (V_{i-1} - V_i)], \quad 0 < \alpha < 0.5, \quad (11.3) $$

where $\Delta V_i$ is the change in $V_i$ from one time step to the next (i.e., $V_i(t+1) - V_i(t)$) and $i+1$ and $i-1$ are the adjacent units. $\alpha$ is a diffusion-rate parameter: larger values of $\alpha$ produce faster diffusion. **Equation 11.3** is a discrete-time version of the standard diffusion equation.

The effects of stimuli and of reinforcement and nonreinforcement are incorporated via two assumptions:

1. **Activity.** A unit is either active or inactive, depending on whether or not its stimulus is present and attended to. If its stimulus is absent, the unit is inactive and the expectation variable changes according to diffusion (Equation 11.3). If its stimulus is present, the unit is active and the expectation variable is set to a fixed value, which is determined by whether reinforcement or nonreinforcement occurs during that time step. (Note that this assumption explicitly incorporates reinforcement into the generalization model, a feature lacking in the simple version discussed earlier.)

To accommodate the free-operant situation and nonspatial stimuli, we assume that a unit is active only when two conditions are fulfilled: the appropriate stimulus is present, and an oper-
ant response is made to it. If the appropriate stimulus is absent, or if it is present but not responded to (not attended to), then the unit is treated as inactive.

2. Reinforcement. Reinforcement sets $V_i(t)$ for the active unit to a value directly related to reinforcement magnitude, $S(t)$. ($S(t) = 1$ in all our simulations). If reinforcement occurs, reinforcement expectation equals reinforcement value; if reinforcement is absent, reinforcement expectation equals zero.

These two assumptions allow the expectation surface to change in real time as a function of the temporal patterns of reinforcement and nonreinforcement in the presence of different stimuli. When the scheme is generalized to two dimensions, I introduce an action rule that allows movement to be guided by the expectation surface. But first, let’s look at the diffusion process itself.

The process is illustrated in Figure 11.5, which shows 20 units, corresponding to 20 values on a stimulus dimension (inset). At time 0 the expectation variable is everywhere zero ($V_i = 0$ for all $i$) and no unit is active. Suppose that at time $t = 1$, the stimulus appropriate to Unit 10 is presented so that Unit 10 is active; suppose also that reinforcement is presented at $t = 1, 5, 9, \ldots$ (i.e., every $4^{th}$ time step) until $t = 24$ and then omitted thereafter, and the stimulus turned off. The activation profile builds to an asymptotic form with successive reinforcements ($V_{10}$, the expectation of the active unit, is zero during training [light lines] just because this figure shows the state of the network only on nonreinforced time steps; on reinforced time steps, $V_{10} = 1$). When reinforcement ceases and no unit is active, diffusion soon creates a typical generalization gradient with a maximum at $S^+$ (Stimulus 10: heavy line). This gradient slowly collapses into a Gaussian form (intermediate line). If the gradient is sampled by presenting stimuli along the continuum, a roughly similar profile of responding can be generated. This one-dimensional process is qualitatively consistent with well-established properties of stimulus control in reinforcement learning: generalization, peak shift and the effects of reinforcement in the presence of multiple stimuli. It also shows a simple kind of spatial extrapolation.

**Generalization.** Stimulus generalization reflects two kinds of processes: perceptual processes that determine the discriminability of one physical stimulus from another; and reinforcement-expectancy (similarity) processes that determine how the animal should respond to stimuli more or less discriminable from the reinforced stimulus. In the diffusion model, the perceptual processes are represented by a single dimension (line of units) along which stimuli are varied. The diffusion process models the expectation that guides action.

Stimulus generalization in animals is demonstrated in two-phase reinforcement-learning experiments (Guttman & Kalish, 1956) like the procedure used to make Figure 11.5. In the first phase, an animal such as a pigeon is trained with an S+ such as a wavelength of 550 nm, in the presence of which responses intermittently produce food reinforcement. In the second phase, no food is delivered and a counterbalanced series of stimuli more or less similar to S+ is presented. The usual finding is that response rate is maximal in the presence of S+ and falls off smoothly in the presence of stimuli increasingly different from S+ – yielding a profile of response rate vs.
wavelength (say), a generalization gradient, resembling the “extinction” curve in Figure 11.5.

**Peak Shift.** One of the more reliable phenomena of generalization involves not one but two stimuli: an S+, as before, and another stimulus, S-, a little different from S+, which is alternated with S+. Responses are never reinforced in the presence of S-. When generalization is tested in the second phase, the result is often a gradient with a peak shifted away from S+ in a direction opposite to S-(peak shift; Hanson, 1959). If S+ is 550 nm and S- is 560 nm, the new peak might be at 540 nm, for example. The magnitude of peak shift depends on the separation between S+ and S-: the closer they are, the larger the shift. Peak shift is usually thought to require some kind of inhibitory process.

**Figure 11.6** shows a simple simulation of peak shift with the diffusion model. S+ (Stimulus 10) and S- (Stimulus 11) are alternated at 4-time-step intervals for a total of three alternations (24 time steps). Two time steps after the end of training, the gradient (heavy line) shows a peak shifted from S+ away from S-. It is also possible to show that peak shift declines and then disappears in the model as S- is displaced further from S+. Notice that the shift here is achieved without assuming that S- is inhibitory (inhibition seems to be necessary to explain behavioral contrast [Reynolds, 1961], but this demonstration shows that the inhibitory assumption is not necessary for peak shift).

The diffusion model is also compatible with Terrace’s (1964) classic observation that “errorless” discrimination training does not produce peak shift. In the model, peak shift occurs because reinforcement expectation is zeroed during each time step when an unreinforced response occurs in the presence of S-. But in errorless discrimination training, no such responses occur, so that the postdiscrimination gradient has the same general form as the single-stimulus gradient.

The diffusion model also produces qualitative results similar to data from experiments with more than one S+ or discrimination studies with multiple S-s (Kalish & Guttman, 1959; Hanson, 1961).

These simulations assume that the rate of diffusion is constant, which implies that if generalization is tested at different times after the last reinforcement (all else remaining the same) the height of the gradient should decrease and breadth of generalization should increase. Both these effects were observed in the original study by Guttman and Kalish (1956). The authors made little of the increasing gradient breadth, even though the opposite result was expected, but Riley (1968) has pointed out that “the decrement in stimulus generalization appears to be present from the first three trials...The gradients appear to flatten with the passage of trials.” (Pp. 21-21.)

Nevertheless, the flattening effect is not always observed and in any case, individual generalization gradients are evidently quite variable in form during extinction. Evidence from memory studies with both humans and animals suggests that older memories change more slowly than newer ones (Chapter 6; see also Rubin & Wenzel, 1996; Staddon, 1998). Hence, a more realistic (but also more complex) version of the model might include an assumption about the rate of diffusion as a function of the “age” of the diffusing effect. Such an assumption would make different predictions about the change in gradient form with time, but would make much the same qualitative predictions about peak shift and multi-stimulus effects, as the simple model.
None of the qualitative predictions depend on this aspect of the model.

**Spatial Extrapolation.** The same process that produces peak shift can also produce a sort of spatial extrapolation. In honeybee foraging, for example “During training to an artificial food source, there comes a point at which at least some of the bees begin to ‘catch on’ that the experimenter is systematically moving the food farther and farther away, and von Frisch recalls instances in which the trained foragers began to anticipate subsequent moves and to wait for the feed at the presumptive new location.” (Gould & Gould, 1982, p. 281). Gould and Gould term this extrapolation “slightly eerie.” I have not been able to find a detailed account of this phenomenon, but Figure 11.7 shows how it might arise from the diffusion model. The left panel shows the spatiotemporal pattern of reinforced and unreinforced locations experienced by bees in a study like von Frisch’s. On Day 1 they are fed at the closest source (first filled circle in the left panel, Site 1). On the second day, no food is available at the first site (gray circle in the second row) but after a delay shorter than a day they find food at the more distant source (filled circle in the third row, Site 2). On the third day, the expectation profile shows the familiar peak shift, peaking not at Site 2, but at a more distal point, Site 3, which has never been reinforced. Extrapolation occurs if the bees are successively fed at any number of more and more distant sites, arranged in a line. Obviously, the exact form of expectation profile at different times depends on

![Figure 11.7 Spatial extrapolation. Left panel: Spatiotemporal pattern of reinforcement and nonreinforcement in an experiment by von Frisch in which honey bees were fed each day at a point farther and farther from the hive. Filled circles: reinforced sites; gray circles: visited, unreinforced sites; empty circles: unvisited sites. Right panel: Light lines: Expectation profile vs. distance after reinforced visits to Site 1 (Stimulus 10, Day 1) and Site 2 (Stimulus 11, Day 2). Heavy line: Expectation profile on Test (Day 3); dashed lines: expectation profile after additional delay. See text for details.](image)

the details of the bees’ search pattern, the time and distance between food sites. Unfortunately, the necessary data on actual bee behavior are not available for comparison. Nevertheless, Figure 11.7 illustrates the fact that a simple form of spatial extrapolation is a robust property of this kind of diffusion model.

**Two-Dimensional Model: Spatial Navigation**

Spatial navigation (a slightly preferable term to orientation) has two logical parts, knowledge and action. Knowledge is the mapping between states of the organism and location in space — Tolman’s cognitive map. The map may be more or less complete, more or less accurate and fine grained. A minimal cognitive map must at least have a defined state (termed a node or unit in our model) for each spatial position (within some spatial resolution), that is, a many-
Figure 11.8

one mapping of spatial locations on to map nodes. For every location there should be one and only one node, although in a coarse-grained map one node may correspond to several neighboring physical locations. Richer maps will have in addition information about adjacency (what is next to what), connectivity (what can be reached from what), and distance (how far places are from each other).

We assume the existence of a navigation process sufficient to locate the animal accurately, that is, to change the active unit in the cognitive map as the model organism changes its position (in a spatial map, only one unit can be active at a time). With the sole exception of learning about barriers (see below), we are not concerned with how the animal knows where it is: any process — landmark learning, dead-reckoning (path-integration), viewpoint memory, GPS etc. — that can provide the required mapping of position on to state will do (see Gallistel, 1990, and Schmajuk & Blair, 1993, for discussions of these issues).

Even an unconnected map permits navigation, because it allows the organism to associate actions with states (locations): the organism can learn to turn left when in State A, right when in State B, and so on. But an unconnected map cannot permit extrapolation or insight (see examples below) because it contains no information about adjacency and because no information can pass from one state to another. Real cognitive maps are at least partially connected. A 4-neighbor fully connected map is shown on the left of Figure 11.8. The other two panels in the figure show different types of map connectivity.

Unit Properties. Each map unit is characterized by three things: it is active or inactive, in the sense described earlier, it has a level of the expectation variable, $V_i$, and it is connected (linked) to one or more units (termed neighbors) representing adjacent spatial locations.

Diffusion Rule. Equation 1 described the diffusion process along a single dimension. At each step the expectation of each unit moves toward the average of its neighbors — one on either side of the current unit. When the map has two dimensions the form of the equation is the same, but the number of neighbors increases. Thus, the generalized diffusion equation is:

$$\Delta V_i = \frac{\alpha}{N_i} \sum_{j \neq i} k_{ij} (V_j - V_i), \quad 0 < \alpha < 1,$$

(11.4)

where $\alpha$ is the diffusion-rate parameter, $N_i$ is the number of units connected to unit $i$ and $k_{ij}$ is a parameter that is equal to 1 when unit $j$ is connected to unit $i$ and 0 otherwise (For simplicity, I denote each unit [node] in the map by a single subscript, even though the map is 2-dimensional).

The only difference between Equations 11.3 and 11.4 is the number of neighbors, $N_i$, associated with each unit, which depends on the type of map and the location of the unit in the map. In the simulations that follow we use an 8-neighbor lattice ($N = 8$), except at the edges of the simulated surface. Units on an edge have only five neighbors; units at a corner have only three neighbors.

The activation and reinforcement assumptions are the same as in the one-dimensional case. For the 2-dimensional case, we need to add an assumption about barriers, as well as an action rule.

Diffusion Control. Diffusion need not occur all the time. For example, it need not occur when no new information has been received — if the organism is immobile, for example. In the spatial simulations described in the next chapter, the diffusion surface is typically updated when a goal is reached, but some updating after, or even during, each movement is also desirable. Diffusion need not occur always at the same rate. Setting diffusion rate according to the age of the initiating event, for example, is an alternative to switching the process on and off. Clearly, diffu-
sion with absorbing barriers should not continue indefinitely, because expectation will eventually approach zero everywhere. When and how diffusion rate should vary is an aspect of the diffusion model that is not yet settled.

**Barriers.** Barriers reduce the number of neighbors and this reduction affects both diffusion and the action rule: diffusion cannot occur between units separated by a barrier and the model organism cannot move between them. Thus, if the action rule requires movement that would take the model organism from unit $i$ to unit $j$ and movement is blocked by a barrier, then parameter $\kappa_{ij}$ is set to zero and $N_i$ is reduced by one. During future time steps, no diffusion can take place between units $i$ and $j$ and the action rule will not initiate movement from one to the other. This is the only kind of map formation (as opposed to map reading) incorporated in the model.

**Action Rule.** For a map to be useful, it must be combined with an action rule that tells the organism what to do at any map location: stay, move forward or back, left or right, etc. Our action rule is very simple: at each time step, the system moves to the adjacent (connected) unit with the highest expectation ($V$-value). If the current unit has a higher expectation than its neighbors, the organism remains where it is; otherwise, it moves to the highest-expectation neighbor. This rule is variously termed hill-climbing or gradient descent, depending on the sign assigned to “good” events.

This system will never become trapped in a local maximum, because of the reinforcement rule and the dynamic nature of the expectation surface. If no reinforcement is available at a given point, $V_i$ is set to zero, which ensures that adjacent points will generally have higher $V$ values. If a reinforcer is present, it is consumed, at which point, $V_i$ again equals 0 and the action rule ensures movement during the next time step.

Notice that this model is both deterministic and local: diffusion flow and the direction of movement from any unit, are determined only with respect to immediate neighbors. The model involves no overview of the problem (or expectation landscape) as a whole.

**Initial Conditions.** In all our simulations, the initial conditions were a uniform low expectation level at each node in the map. Under these conditions, the system will actively explore its environment, zeroing expectation at each point it passes through.

The next chapter illustrates how these assumptions work to simulate real spatial navigation in animals.
Chapter 12

SPATIAL SEARCH

The diffusion-generalization model defines a route to be followed within the map. The route is controlled by an expectation surface that is altered dynamically both by the passage of time and by the model organism’s experience with reinforcement and nonreinforcement. This chapter describes how the behavior of the model matches qualitatively the behavior of searching animals in both open-field and constrained conditions (Reid & Staddon, 1998).

Figure 12.1. How the expectation surface changes as the animat moves in area-restricted search. Four sample surfaces are shown, at 15-step intervals. The point on the animat path corresponding to each surface is shown by the arrows. The animat position corresponds to the small filled square on each surface. The goal is reached after 30 steps; 60 steps are shown.
Open-Field Foraging: Area-Restricted Search.

If a rat is allowed to find buried food at a particular spot in a large enclosure, removed and then returned after a delay, it usually begins digging close to the reinforced spot. If it fails to find food, it searches the immediately surrounding area before beginning a more general, unlocalized search. Area-restricted search has been observed in a variety of species, including shrews (Pierce, 1987), pigeons (Cheng, 1989), and even insect larvae and ants (Banks, 1954; Bond, 1980; Harkness & Maroudas, 1985; Nakamuta, 1982; Wehner, & Srinivasan, 1981).

The 2-dimensional diffusion model produces area-restricted movement patterns when the simulated organism (the term animat is sometimes used) is reintroduced to the enclosure after finding reinforcement at a particular spot. The process works like this. In a new situation, the model assumes initially a low, uniform level of reinforcement expectation at every unit in the map. This small a priori expectation causes the animat to move around the enclosure in an exploratory way as successive failures to find food "zero" expectation at each spot. Once the animat finds reinforcement at a particular spot, expectation for the active unit is set to unity and the animat is immediately removed from the enclosure (i.e., all units are made inactive). Reinforcement expectation at that position immediately begins to diffuse according to Equation 11.4 at a rate determined by $\alpha$. Diffusion occurs for a number of time steps proportional to the number of units in the map — so that information about reinforcer is transmitted to the map borders. Diffusion does not continue indefinitely, else all information would be lost. (I discuss later how diffusion rate should be controlled.) Figure 12.1 shows the pattern of movement produced when the animat is subsequently reintroduced to the empty (no-food) enclosure. The figure shows both the search path and the expectation surface at four different positions along the path.

The hill-climbing action rule moves the organism directly from the START to the peak of the expectation landscape. This movement has the appearance of goal-direction (i.e., the animat moves straight to the goal) but in fact only local processes are involved: the animat cannot sense the peak of the expectation landscape from a distance. At each unreinforced step, the value of the reinforcement expectation for the active unit is set to zero. The diffusion process (updated for a few iterations at each step) allows reinforcement expectation to "flow" back gradually into the units the animat has already visited (like tracks made by walking through a viscous puddle). The amount of flow back into previously visited spots depends upon the amount of reinforcement expectation in the neighboring units and the time elapsed — the higher the expectation, the longer the time elapsed, the more flow, hence the more likely the organism will revisit the spot. The dynamics

Figure 12.2 Top panel: Distance from the origin ("nest") of a single searching ant at 2-m intervals as a function of total path length, in m (from Wehner & Srinivasan, 1981, Fig. 8b). Bottom panel: Distance from goal as a function of time steps of a simulated organism searching according to the diffusion model. The model was allowed to iterate for 20 steps after reward was presented at the origin before the animal was allowed to begin searching. Diffusion rate, $\alpha = 0.1$. Very similar patterns are found with $\alpha$ values ranging from .02 to .1 and initial iterations from 20 to 50 (from Reid & Staddon, 1998, Figure 7).
of the expectation landscape act like a working memory that shows spontaneous recovery, which I discussed in another context in Chapter 6.

With each unreinforced step, reinforcement expectation is zeroed at that point. As more and more steps are made near the reinforced spot, the overall level of reinforcement expectation in the area declines, and the animat gradually moves farther away to areas in which expectation has not been depleted. This gradually expanding search pattern closely resembles descriptions of "focal" search followed by more "general" search in rats (e.g., Timberlake & Lucas, 1989).

Area-restricted search is centered on the reinforced location. Movement in all directions from that spot is equally likely, as it is with actual animals (Cheng, 1992; Wehner & Srinivasan, 1981). Even though the model has no stochastic component, the search path appears erratic and frequently intersects itself. It is chaotic, in the strict sense. This general pattern is not dependent on the value of the diffusion-rate parameter $\alpha$.

**Area-restricted Search: Fine Structure.** Desert ants of the genus *Cataglyphis* find their nest — a one-centimeter hole in the ground in an almost featureless environment — by integrating the position change produced by each step (*path integration*), with nest-site as the origin. (Unlike many other ant species, they do not use odor trails.) Wehner & Srinivasan (e.g., 1981) have provided beautiful data on the search paths of these animals as they look for the nest. The ants show a pattern that looks like area-restricted search, but has interesting fine structure, which is predicted by the diffusion model.

A typical experiment is as follows: "Individual ants were picked up at a feeding station, mostly at a distance of 20 m from the nest, transferred to a small glass flask and displaced to the test area which was located 300-600 m away from the nest area. There the ant was released and its return run recorded by means of a grid of white lines painted on the hard desert plain." (p. 317) Because these animals’ map is created through path-integration, the ant is completely unaffected by passive displacement. It sets off in the same compass direction it would have, had it never been moved. But when it reaches the point where the nest should be, the ant turns and begins to search in a fashion that closely resembles the path in Figure 12.1. In other experiments, a returning ant was taken from the mouth of the nest: in this case, the animal begins to search at once at the new location, since its path-integration system tells the ant it has reached the nest.

Although the search pattern looks pretty irregular, there is in fact structure in it: “Generally speaking, the ant performs a number of loops of ever-increasing diameter, pointing in different azimuthal directions. This
system of loops is centred around the origin [i.e., the nest location] to which the ant repeatedly returns during searching...Even when the ants, after searching times of more than half an hour, have departed from the origin for more than 50 m, they may return close to the origin before starting for another loop in another direction.” (P. 319) Commenting on this looping pattern, the authors note that “No navigational mechanism is reset when the ant has arrived at (or near to) the origin...the return to the centre must be regarded as an intrinsic feature of the search strategy itself.” (P. 319)

One way to summarize this looping pattern is to plot the ant’s distance from the origin (nest) as a function of time (or, equivalently, path length, since the ants move at a more or less constant speed). A typical example is shown in the top panel (“Data”) of Figure 12.2, which is taken from Wehner and Srinivasan (1981), Figure 8b. The figure shows distance from the origin on the y-axis vs. path length (in m) on the x-axis. The division between an initial “focal” search — small excursions away from the goal (when path length is less than about 100 m) — and “general” search, when the animal drifts farther and farther away from the goal in between return swings, is clearly apparent in the distance vs. path-length record. But the most striking feature is the oscillations of increasing amplitude late in the search as the animal swings in to and away from the goal. A typical simulation with the diffusion model is shown below the data in Figure 12.2 and the general pattern is very similar: focal search followed by wide return swings farther and farther away from the goal. This same pattern is found with a range of $\alpha$ values.

Wehner and Srinivasan explain why this periodicity is adaptive: “there can never be complete certainty...that a given region has been fully explored. All that can be said is that the probability of encountering unexplored areas within the region decreases as the time spent in searching the region increases...Another way of expressing this idea is to say that the probability of finding the nest within a given region decreases as the time spent searching that region increases...The entire searching strategy [and the periodical return to the origin] derives from, and is based upon, this one simple concept (p. 326).”

Wehner and Srinivasan propose a stochastic optimal-search model that embodies this idea. Their model matches the data pretty well (see their Figures 15, 16, 19 & 20), though not, quite as well as the diffusion model (the oscillations of their model are too regular). But the diffusion model perfectly captures the core idea: that search is centered on an origin; and that the probability of visiting a region is inversely related to time spent in the region (extinction) but positively related to time since the region was last visited (spontaneous recovery). It is these two factors in combination that lead to periodic returns to the origin: At first, a sector (termed “azimuthal direction” by Wehner and Srinivasan) near the origin is extensively explored, because
expectation is highest there. As expectation in a sector near the origin is depleted, the animal moves away. But once it moves away, expectation can flow back from yet-unexplored sectors close to the origin, and eventually this increase in expectation causes a return to the origin, which allows another sector to be depleted. This process is repeated, leading each time to a looping back to the origin, until the whole central area is depleted and the search moves outwards.

Open-Field Foraging: Multiple Goals. Area-restricted search is best seen with a single goal, but real foraging usually involves multiple food sites. One of the few multiple-goal studies that measured actual tracks is a simple experiment with a European badger by Mellgren and Roper (1986). For six days, the experimenters placed shelled peanuts at several fixed positions in an open grassy field and recorded the movement patterns of the badger as it entered the area at night, located and consumed the food, and left the area. Each day, the peanuts were placed in the same fixed positions, shown in the top panel of Figure 12.3. The second panel shows the track of the badger on the sixth night as it entered the study area, located and ate the peanuts, and finally left the area (data from the preceding nights were not presented).

The bottom panel of Figure 12.3 shows a typical simulation result. The simulated rewards are distributed in the same spatial pattern as in the Mellgren and Roper experiment (top panel of Figure 12.3). As in the prior simulations, the animat was allowed to locate the rewards and was then immediately removed from the area. In Mellgren and Roper’s experiment, sessions were separated by approximately 24 hours, and the badger typically spent about 15 minutes in the study area collecting the peanuts. We simulated the 24-hour period by iterating the diffusion equation 300 times with a low diffusion rate ($\alpha = 0.001$) before returning the organism to the area. We stopped the simulation when the organism left the area (i.e., moved as far away from the reward as the point of introduction). As before, each “step” of the organism iterated the diffusion equation once. Since movement patterns of actual animals are not identical each night in this procedure, the movement generated by the model cannot be expected to produce patterns identical to the data. Moreover, the behavior of the diffusion model is very dependent on initial conditions: very small changes in starting point (or in diffusion rate) can produce very different search paths – the simulation may cross between Areas A and C two or three times, rather than just once, for example. But the general pattern will always be the same, showing a complexity and frequency of direction reversals and path recrossings that resembles the badger data.

Short Cuts. The badger in the Mellgren and Roper experiment, as well the animat, crossed from one food patch directly to the other, without returning to its starting point. Not surprising, perhaps, but this result does suggest that the diffusion model has the capacity to take short cuts, i.e., to take a novel, shorter path after experience with a roundabout path to a goal. Chapuis (1987) did a very simple short-cut experiment with dogs. Imagine three open-field locations arranged as in each panel of Figure 12.4: A, B, and the dog, arranged in a triangle, with the dog as starting point. A is closer to the dog than B: distance dog-A < dog-B. A hungry dog on a leash was led from the starting point to Location A where food had been placed, and then led back to the starting location without being allowed to eat the food (Path start-A-start). This procedure was repeated for another, more distant, location, B (Path start-B-start). Path start-A-B is obviously shorter than the sum of the two training paths, start-A-start-B. However, going straight to B after eating the food at A involved taking a short cut, Path A-B, over ground previously unvisited by the dogs. Most dogs tested nevertheless took the short cut (dashed line) when allowed to run free from the starting point.

104 Note that the basic diffusion model is not context-sensitive and has no long-term memory. Without additional assumptions, it cannot account for differential effects in and out of the experimental apparatus or for long-term retention. This slow-diffusion-outside-the-apparatus assumption is a way to accommodate the idea that memory changes more slowly when the animal is outside the learning context than when it is inside.
The diffusion model works like this in the Chapuis experiment. The simulated organism first follows the steeper gradient to the closer patch, A. As it travels, expectation is zeroed at each step, so that once the reward at A is consumed, and that unit is also zeroed, the steepest gradient will be in the direction of the unvisited patch, B, more or less along the A-B line. Given an accurate map, the simulated organism, like Chapuis’ dogs, perhaps after some area-restricted search at A, will go straight to B rather than retracing its path to the start. The model shows spatial “insight.”

**Shortest-distance vs. minimum-delay paths.**

Mammals and birds usually act in a way that approximately minimizes delay in getting food. In operant studies of explicit choice — on concurrent schedules, for example — pigeons, rats and several other species always choose the alternative with the shortest delay to food reinforcement. Moreover, later delays tend to be discounted. For example, given a choice between one key signaling a fixed-interval (FI) schedule with a 20-s delay to a single reinforcement (followed by a return to the choice situation) and another key signaling an FI 15 s, animals will prefer the FI 15. The shorter the first-link schedule (VI T), the greater the preference (see Figure 12.5). If a second reinforcement, with a shorter delay, is added to one second link (Figure 12.6), pigeons may still prefer the side with the shorter first link, even if it has the lower overall reinforcement rate. But if additional closely spaced reinforcers are added, preference will eventually shift. In other words, delay is critical, but remote delays are somewhat discounted (see Williams, 1988, for a review of this literature). These same principles seem to apply to foraging in the open field.

When foraging for multiple feeding sites, animals visit all the sites, but they don’t always follow the shortest-distance path — they are not perfect “traveling salesmen.” (Cramer & Gallistel, 1997; Wehner, Lehrer, & Harvey, 1996). For example, consider a distribution of food sites like that in Figure 12.7: one source close to the starting point, a cluster of sources further way, in a different direction. The traveling salesman (shortest-path) solution is obvious, shown by the line (path 1,2,3,4,5). Yet real animals are likely to go for the cluster first (sites 2,3,4,5), even though it is further away, because doing so minimizes total discounted food delay.
Figure 12.8 shows an example from an actual experiment with vervet monkeys (Cramer & Gallistel, 1997) in which the distribution of goals was set up as an explicit test of the idea that monkeys treat foraging as a traveling-salesman problem, i.e., that they visit food sites in shortest-path order. There are two target clusters: sites 1, 2, 3, 4 and sites 5 and 6 (Figure 12.8, left panel). All the monkeys in the experiment chose first to visit the cluster with the greater number of food sites (1, 2, 3, 4), even though sites 1 and 6 were equidistant from the start. A typical path was 1-2-3-4-5-6.

The right panel in Figure 12.8 shows the path followed by the diffusion animat foraging for the six goals. In the Gallistel and Cramer experiment, sites 1 and 6 were equally close to the starting point, but in our simulation, Site 6 is actually a bit closer to the start than Site 2, which is part of the 4-site cluster. Even so, site 2 was visited first.

The diffusion surface was created using 3000 iterations. The animat started at point S and visited the targets in the order 2, 3, 1, 4, 5 and 6. The travelling-salesman path is less than 100 movements; in this simulation, the animat used 323 movements to visit all sites.

Figure 12.9 Diffusion-surface updating. Top row: Three food sites and the start position (bottom left) at three stages of search. Line shows the search path. Bottom row: Expectation surface initially and after locating sites 1 and 2.

Diffusion Control. This behavior is similar to that shown by the vervet monkeys — the large cluster is visited first — but with more area-restricted searching. The reason for the additional searching is that even if the diffusion surface is zeroed at the current animat position, at the next iteration, expectation will be partially restored (spontaneous recovery) owing to the contribution of neighbors that are still above zero. There are several ways to reduce wandering around a depleted site. The simplest is to update the diffusion surface after each site is visited.

105 These simulations were done with the assistance of Mircea Chelaru.
The way this works is shown in Figure 12.9. Initially, each food site corresponds to a peak in the expectation surface. But after the first site is found, it is held at zero, the others at one, and several diffusion steps are carried out, until the surface stabilizes, as shown in the center panel: two remaining peaks. Then the second site is found, and the process is repeated.

This surface updating seems arbitrary, but it corresponds to the obvious adaptive fact that the surface should reflect the degree of uncertainty the organism has about the location of goals. If the organism is completely certain (based on prior learning or, the human case, verbal instructions), then it makes sense to completely update the surface after each goal-find. But if the environment is new, and the organism has just found food, then there may be food in the immediate vicinity. The surface should only be slightly modified — and area-restricted search is the result. In other words, it makes sense that diffusion rate should be under the control of higher-order associative processes — at least in mammals (probably not in ants). When the environment is familiar (as in the vervet experiment), complete updating after every “find” makes sense.

Figure 12.10 shows the path followed by the animat (in the vervet experiment simulation) when the diffusion surface is updated after each food site is visited. The order of visits is 2-3-4-1-5-6 (left panel of Figure 12.10). Updating reduces wandering. The animat path is not the shortest possible, but the discrepancies are because the animat map is a lattice approximation.

Although the diffusion model has some of the properties of working memory it lacks most of the properties of long-term or reference memory. For example, the model deals with nonreward by zeroing the expectation value of the active unit. If diffusion rate is constant, this means that after nonreward, the only difference between a unit that has been frequently reinforced in the past and one that has been rarely reinforced will be the level of expectation in neighboring units. But if reward resets expectation to a fixed value, rather than cumulating, there may be little or no difference in the immediate environments of previously rewarded units with these two different histories. (Cumulation, in turn, poses difficulties for prediction of simple search and exploration.) Hence, without the addition of diffusion-control processes, the model cannot accommodate histories that have very different effects on future behavior. Real animals will relearn about a spot with a long history of reward faster than they relearn about a spot that has been rewarded only once, for example. This problem is an example of the stability-plasticity dilemma (cf. Chapter 6): when to learn something new vs. when to retain something old. Higher-order diffusion-control processes are necessary for a comprehensive account.

Cramer and Gallistel describe one experiment where their monkeys tended to follow the minimum-distance path. The food sites were in the diamond pattern shown in Figure 12.11. When the animal begins searching, no food is present at the start (Site 1); all others are baited. The animals usually chose either route 1234 or 1324, in other words, at each point they chose the
closest next goal, as would be predicted by the diffusion model. But in another condition, food was placed at site 1 after the animal reached one of the other sites. Now either of the previously preferred routes (ending at site 4) would much delay the animal’s return to the starting point and, as the minimum-delay idea would predict, they tended to choose the diamond route: 1243 or 1342. The fact that adding food at site 1 would displace the next choice, at sites 2 or 3, from either sites 3 or 2 to the more remote site 4 is not consistent with the diffusion model. It is consistent with the idea that the monkeys sampled both the zigzag (e.g., 1234) and diamond (e.g., 1243) routes and always chose the minimum-delay route — as I suggested when I compared these experiments to the concurrent-chain operant schedule earlier. The implication is that both the diffusion process (necessary to explain cluster choice, as in Figure 12.8) and route-preference (necessary to explain minimum-distance choice, as in Figure 12.11) may operate at different times, under the control of higher-order processes. Absent a general theory that can account for all the data with a unified process, this conclusion, though wimpy, may be true.

The diffusion surface in the vervet experiment was “set” by pre-exposing the animals to baited food sites. The resolution of the surface is therefore subject to memory limitations — which vary substantially between species. Some birds, such as Clark’s nutcracker (Nucifraga columbiana) bury food in thousands of caches and seem to remember them more or less individually, for example (Shettleworth, 1995). In experiments like the vervet study, chimpanzees are able to remember 18 locations, as opposed to the vervet’s six (Menzel, 1973). Better spatial memory implies more nodes — higher resolution — in the diffusion surface.

**Detour Problems.** If a barrier is introduced into a familiar environment, most mammals and birds soon learn to maneuver around it. This is the classic umweg (detour) problem used by the Gestalt psychologists and others to demonstrate spatial ‘insight’. Simulation of open-field behavior assumes a fixed cognitive map. Learning about barriers obviously requires modification of the map. I assume that encountering a barrier breaks the link between two adjacent units: parameter $k_{ij}$ in Equation 11.4 is set equal to zero and $N_i$ is reduced if a barrier is encountered when the animat attempts to move from node $i$ to node $j$. Expectation will not flow directly between the nodes and the animat will not attempt to move from one to the other. Encountering a barrier does not affect the animat’s behavior directly, it only changes the map. There is an indirect effect, because the changed map changes patterns of diffusion and, therefore, changes the form of the expectation surface.

There are two classic types of detour problem. One involves the placement of a U-shaped barrier between the subject and the reward in a relatively open field (e.g., Hull, 1938; Lewin, 1933; Tolman, 1932). The other involves the use of removable partitions within an enclosure (e.g., Dennis, 1929).

U-barrier problems entail placement of a U-shaped barrier between the subject and a reward. Since the direct route to the reward is blocked, an indirect route must be found if the reward is to be obtained. Inexperienced or less cognitively elite subjects, such as rats or young children, first attempt the direct route, only later taking an indirect route that approximately minimizes the (indirect) distance between the starting location and reward. Experienced subjects take the indirect route immediately without attempting to go through the barrier. The model implies that the difference between smart and less smart subjects is in their ability to change their maps, not in the level of “cognitive processing” after they’ve learned the new map. Once the map is changed, the diffusion model solves the problem.

We (Reid & Staddon, 1998) began simulations of the U-detour problem with the assumption that all adjacent units are fully connected, including those on each side of the barrier: that is, there was no a priori knowledge of the barrier. Every unit was initialized with the same low value of reward expectation. In a free field this initial condition produces an exploratory pattern of movement. As before, the diffusion equation iterated once for each step taken by the simulated organism. This process brought the organism into contact with the barrier at various points.
Each time the barrier prevented the organism from moving to a neighboring spot (i.e., a neighboring unit), the two units were declared nonadjacent (\(k_{ij}\) set to 0, \(N_i\) reduced by one). The simulation continued until the barrier had been encountered at each point. Subsequently, reinforcement was placed at one of two goal locations, the simulated organism at one of two starting locations on the other side of the barrier. The simulation continued until the reward was found, at which point the organism was immediately removed from the area (all units were deactivated). The diffusion equation was iterated for a further 30 steps, and the system was then placed at one of the two starting locations on the other side of the barrier from the reward. The simulation was then iterated until the reward was found.

In each case, the simulations produce a movement path which approximately minimizes the distance between the starting position and the reward location, given the constraints imposed by the lattice map structure. These movement patterns result from the diffusion of reward expectation around the barrier. The patterns do not depend upon the value of the diffusion rate parameter, \(\alpha\). The movement path closely resembles reports in the Hull, Lewin and Tolman studies cited above as well as in newer studies (e.g., Chapuis, 1987) with this task.

The second type of detour problem involves the use of removable partitions in an enclosure. In an early report, Dennis (1929) described the paths taken by blind\(^{106}\) rats in a rectangular enclosure containing two removable partitions. As individual rats moved about in the enclosure, their paths were traced by means of a camera lucida. The simulations began with all adjacent units connected, both inside and outside the enclosure. That is, there was no a priori knowledge of partitions or walls. Every unit was initialized with a low expectation value to produce an exploratory pattern of movement. This exploration eventually brought the organism into contact with the walls of the enclosure and any internal partitions. Each time a barrier prevented the organism from moving to a neighboring unit, the two units were declared nonadjacent.

The diffusion process was iterated once for each step or attempted (i.e., barrier-blocked) step. The simulation continued until all the barriers had been identified. Then the animat was removed from the enclosure and placed in the start box. When the animat’s continued exploration eventually led it to the goal box, it was immediately reinforced for one time step and moved back to the start box for the next trial, which began immediately.

At the beginning of the simulation, the animat explored the enclosure. Movement was mostly restricted to the walls and partitions. The animat moved into the open areas only after all the partitions and walls had been traversed at least once. This finding strongly resembles so-called thigmotaxis: rats exploring a large novel enclosure at first restrict their movement to the walls and only much later venture out into the open. After the animat had explored the enclosure and obtained reward in the goal box, the patterns of movement closely resembled the patterns obtained by Dennis (1929). When only one partition was in place, the simulation produced movement that approximately minimized the number of steps from the start box to the goal box. The animat achieved a similar minimum-distance solution when a second partition was added. The main difference between the movement of the animat and actual rats is the occasional abrupt changes in path direction because of the coarse grain of the map: the finer the map, the closer is the predicted movement pattern to the minimum-distance solution.

After Dennis’ rats had been trained to run through the enclosure with one partition in place, the partition was removed and the trials continued. Dennis describes the behavior of the rats this way:

The path which they had followed from 87 to 100 trials gave way immediately to a shorter one.

---

\(^{106}\) The fact that these rats were blind undoubtedly made it difficult for them to form a map, but evidence from later experiments (e.g., Dale, 1982) suggests that once the map is formed, blind and sighted rats behave very similarly.
Of the total records during the three days when the wall was lifted, 64% show some trace of the old habit. In this percentage of cases, that is, the rats went to the right of a straight line between the entrance- and food-boxes . . . . There was a general progression of the paths toward a straight line, but the behavior varied from trial to trial and from rat to rat. (1929, p. 70)

The behavior of the animat resembles Dennis’ description in several respects. On consecutive reinforced trials, the distance traveled from the start box to the goal box is shorter than when the partition had been in place. The simulations also show a similar trace of the prior “habit” — and the simulation produced a similar progression of successive routes towards a straight line.

Dennis argued that the pattern of path changes when the partition was removed shows that spatial learning is not simply based on S-R associations, because the “maze habit” does not show up as identical behavior on successive trials. The simulations show that once the dynamics of generalization are taken into account, an S-R process works pretty well.

Mazes

Mazes have been widely used in the study of spatial learning. Adaptive behavior in mazes has often been used to justify large claims about cognitive processing. For example, Tolman, Ritchie, and Kalish (1946) argued that rats often take the shortest path to the goal, whenever such a path is available (i.e., not prevented by the design of the maze) and that this behavior provides evidence for a type of spatial insight. I have already described how an essentially local model is capable of producing minimum-distance maze paths under many conditions. I show in this section three other examples of intelligent maze behavior that can be simulated by the diffusion model.

Tolman’s “Insight” Mazes. The short-cut problem of Tolman et al. (1946) using the “sunburst” maze is probably the best known example of spatial insight in animals. Rats were given extensive training on the elevated maze shown on the left in Figure 12.12. The rats were first trained to run down the short path from the “start” then across the open field and down the “left-right-right” path to the goal (left-hand maze). In the test phase they began again at the start in the “starburst” arrangement on the right but the straight-ahead path was blocked at point A. The rats usually (but not invariably) chose instead Path P, pointing directly at what used to be the goal, rather than one of the paths on the left corresponding to the left dog-leg in the original maze. Tolman et al. concluded that training produces (in some rats, at least) a “disposition to take the shortest Euclidean Path to the goal.” (1946, p. 20)

The diffusion model easily simulates this result. Given an accurate cognitive map and known goal location outside the maze, the values of expectation outside the open-field area are always higher than values inside, so the animat will persistently explore the edges of the open field. In this way it discovers barriers at the edges of the open field and around the straight-ahead-left-right-right segment that leads to the goal. At this point, there will be a diffusion gradient along the runway with a peak at the goal, so the animat will move smoothly to the goal along the runway in the arrangement on the left in the figure. But when it is transferred to the “sunburst” arrangement on the right, the block at A forces the animat to explore again the edges of the open field, where there

Figure 12.12 Left: Schematic picture of the training maze used by Tolman, Ritchie & Kalish (1946). Right: The “starburst” maze used by Tolman et al. (1946) during testing with a block at A. P is the path pointing at what used to be the goal.
are now open runways. The expectation gradient is always highest in the direction of the goal, so that once the new open alleys are discovered, the simulation will choose Path $P$, the “insightful” solution.

The famous Tolman and Honzik “insight” problem discussed in the previous chapter (Figure 11.4) can obviously be solved in the same way. Given an accurate map and knowledge of a barrier at point $B$, expectation can only diffuse back to the start box only along Path 3, which should therefore be the one chosen.

**Radial-Arm Maze.** The radial-arm maze first studied by David Olton (e.g., Olton & Samuelson, 1976) is a sort of inverse of the classical single-goal maze and produces very different behavior. A nice feature of the diffusion model is that it applies just as easily to the radial-arm maze as to classical mazes even though the observed behavior is so different. Training in the radial-arm maze proceeds in two steps. First the animals are allowed to explore the maze (usually, just once per day) with food at the end of all (usually eight) arms. They explore the maze in an efficient though not necessarily patterned way that ensures that they revisit no arm until all have been explored. This behavior is sometimes termed “paradoxical” since conventional law-of-effect learning would presumably have the rats always returning to the first rewarded goal box. Suggestions about “exploratory tendencies” and “spontaneous alternation” have been offered to account for this apparently puzzling behavior (Olton, Walker, Gage & Johnson, 1977; see review in Gallistel, 1990).

In the second phase of a typical radial-maze experiment, the rat is allowed to enter only some (e.g., four) of the arms. It is then removed from the maze and returned only after a substantial delay, whereupon it enters first just the unvisited arms (Olton & Samuelson, 1976). This result also caused a stir, because the animals choose accurately even if removed from the maze for several hours after entering the first four arms — a delay much longer than that sustainable under conventional delay-of-reinforcement procedures. In fact, the rats’ ability to remember across a long delay is not as puzzling as it may seem, given the lengthy intertrial interval in these experiments as well as other differences between this procedure and conventional delay-of-reinforcement procedures (Lett, 1975; Staddon, 1985). Thus, the biggest puzzle posed by radial-maze experiments is why the animals are so accurate in the first place. Why do they not always return to the first rewarded arm? Why do they explore the other arms without revisiting? Do we need to postulate some kind of special principle or instinct, or can this behavior be derived in a natural way from the processes of spatial learning?

The diffusion model provides a straightforward answer to these questions. Given an accurate cognitive map, behavior in a maze depends solely on the initial conditions (i.e., the expectation surface at the beginning of testing) and the subsequent reinforcement schedule. Given a history of training with a single goal box (as in the Tolman experiments), the expectation surface in the maze will be peaked at the goal which, as we have seen yields goal-directed behavior. But given no training, we may expect a flat expectation surface which yields exploratory behavior — and would yield spontaneous alternation in a T-maze or varied paths in a Dashiell maze (Dashiell, 1930). If exploratory behavior leads to finding reward at the end of the first arm in the radial maze then, after a period of area-restricted search, the simulated organism will backtrack and choose another arm, whose expectation strength must be higher than the arm just extensively (and unsuccessfully) explored. If it finds food again, the process will be repeated, until each unvisited arm has been visited once. The result will be a pattern of search in which the least recently visited arm will normally be highest priority. Moreover, the expectation strength of each arm visited will be reduced relative to the arms unvisited since the previous day, until all have been visited. In a realistic model, speed of decision making will be inversely related to the expectation differences to be compared: the smaller the differences, the longer the time to decide between alternative moves. Such an animat will, like the real animal, hesitate before searching again once all arms have been visited on a given day.
Thus, the diffusion model easily explains the basics of radial-arm-maze performance and does so without invoking any novel principles. At the beginning of testing, the expectation surface in any single-goal maze, such as Tolman’s or the Morris (1981) water “maze,” will have a single peak at the goal. The surface for the radial-arm maze will be flat or have multiple peaks. The difference between behavior in the standard and radial-arm mazes is a consequence of the different initial conditions (diffusion surfaces) at the beginning of testing with the two maze types. It need not reflect some special set of processes unique to rats.

**Constructing the Map**

The diffusion model assumes a map. How is the map made? This may be less difficult than is often assumed. For *Cataglyphis*, an x-y coordinate system is presumably built-in. The animal uses the sun to orient the map with respect to the world, and then path-integration gives it its current position with respect to the nest at the origin.

Mammals use landmark learning, but even this need not involve anything very complex. All that is necessary is a process that maps the space of landmark azimuths (2-dimensional for two landmarks, 3-dimensional for three, etc.) on to a 2-dimensional array of neural units. Schmajuk and Blair (1993) have recently done this using a set of neural-network detectors “tuned” to landmark azimuths and trained by back-propagation reinforcement to generate generalization surfaces with properties very similar to our expectation surface.

Hampson (1990) and Schmajuk and Thieme (1992) have proposed connectionist schemes that also combine map-building and route finding. Instead of diffusion, Hampson uses forward and backward “spreading activation” to create “subgoals,” which guide movement. Like the diffusion model, Hampson’s breaks links between nodes separated by a barrier.

**CONCLUSION**

Once an organism has the capacity to form some kind of internal, “cognitive” map, a very simple route finder is sufficient to guide behavior intelligently. Any map contains a great deal of information. It is perhaps not surprising that a relatively simple map-reading process is sufficient to extract it. I have discussed a route finder that is nothing more than a dynamic formulation of stimulus generalization, a fundamental process in associative learning. Thus, Tolman’s contention that something more than S-R principles is required to account for maze behavior is only partly true. A map is required, but no “insight,” no internal cyclopean eye with an overview of the situation-as-a-whole, is necessary.

The diffusion model is sensitive to only part of the stimulus context. It assumes that in the absence of its stimulus a unit is inactive, but it is also likely that diffusion rate will vary as a function both of context – whether the organism is in or out of the experimental apparatus, for example – and “attention” – what parts of the apparatus are acting on its behavior at any time. Incorporating these effects requires a more comprehensive model.

The diffusion model predicts qualitative properties of search paths. The predictions are qualitative rather than quantitative for several reasons: (a) Because the model is incomplete in the ways just discussed. (b) Because even though the model is entirely deterministic, nonlinearity in the reinforcement assumptions means that the predicted path is highly sensitive to parameter values and to initial conditions (chaotic, in the technical sense). Consequently, the model cannot be expected to predict specific paths. Moreover search paths in real experiments are never reproduced exactly (just as the model implies). For both these reasons – the properties of the model and variability of data – the model cannot be expected to predict specific paths. (c) Because the model lacks control processes that set diffusion rate and number of diffusion steps, prediction of specific changes in search path as a function of time since exposure to a given problem is somewhat problematic.

On the other hand, given parameters (diffusion rate, number of iterations) adequate to predict the search path after exposure to one set of goals, as in Chapuis’ 2-goal experiment, for
example (Figure 12.4), the model should be able to predict search paths in similar experiments with more or differently arranged goals. The assumption about barrier-learning could be tested systematically by varying exposure to barriers and looking at detour paths.

The model implies changes in the expectation surface as a function of time. It therefore predicts reversals in preference for two rewards in an open field as time between reward encounters is varied. For example, in the Chapuis study, the dogs went to the closer food location before they went to the farther location. However, the model suggests that if there is a large difference in time between exposure to the two food locations, the first (closer) location will have a flatter expectation gradient than the gradient for the more recent, but more distant, food. Therefore, animals should go to the more distant food source first.

In general, the model makes definite predictions about changes in preference as a function of the order and times at which rewards are encountered. It is important to emphasize, however, that since the rate of diffusion is a free parameter, the model can only predict a certain pattern of preference change as a function of time. If diffusion is very rapid, for example, then the predicted preference changes may take place over times too short to be practical for behavioral test, and only the long-time preference will be observed.
Chapter 13

TIME, I

Dost thou love life? Then do not squander time, for that is the stuff life is made of. Benjamin Franklin

There is a tide in the affairs of men, Which taken at the flood, leads on to fortune… William Shakespeare

Time is involved in operant behavior in two ways that may, or may not, be related: as a determiner of value and as a cue. It is the denominator in computations of value. Behavioral ecologists, for example, define the profitability of a food source as the ratio of energy content to handling time: E/H. The larger the energy content, or the smaller the handling time, the greater the profitability of the item. The comparable measure in operant conditioning is rate of reinforcement: the ratio of reinforcers obtained to the time taken to get them. In both cases, value is presumed to be directly related to the ratio of some measure of reinforcement to the time period over which the reinforcement was obtained. Time in this sense is “the stuff of life.”

In fact, rate tout court is not an adequate measure of reinforcer value although, other things being equal, an animal will prefer a higher rate of reinforcement to a lower. As I argued in the last chapter, immediacy is also important: a stimulus that signals two reinforcers delayed by 10 and 30 s will be preferred to one that signals delays of 30 and 10, even though the rates of reinforcement are the same for both.

But time is also a sort of stimulus: Activities — like looking for food — are effective at some times and not at others: there is a tide in the affairs of men… Time in this context is a cue to the proper scheduling of different activities. Time as a stimulus is studied as interval timing. I discussed one example, the fixed-interval (FI) schedule, in Chapter 2. If food reinforcement is available only a fixed time after the preceding food, animals wait after each reinforcement, then respond at an accelerating rate as the time for the next reinforcement approaches (Figure 2.7).

How are these two roles for time related? Does the computation of value depend on interval timing? There is no necessary connection. It is easy to devise a choice model, for example, that cannot use time as a stimulus but nevertheless takes rates of reinforcement into account (e.g., Dragoi & Staddon, 1999). And interpretations of choice between variable-interval schedules based on differential interresponse-time reinforcement have had only limited success (Hinson & Staddon, 1983; but see also Cleaveland, 1999). On the other hand, a popular model for interval timing, scalar expectancy theory (SET: Gibbon, 1991), contends that choice and interval timing depend on the same clock-like process. The relation between choice and interval timing is still an open question. But much is known about interval timing itself. I discuss SET in this chapter, and an alternative in the next two.

INTERVAL TIMING

If food delivery is strictly periodic, a hungry pigeon or rat will adapt by waiting an approximately fixed time, proportional to the interfood interval (IFI), after each food delivery before making the operant response (linear waiting: Wynne & Staddon, 1988; also proportion timing when peak rate is the measure). But if food delivery is more or less aperiodic, animals will also adapt, picking up almost any periodicity in the input sequence. Figure 13.1 shows one well-known example. The open-circle line shows response rate as a function of postfood time after long training on an interval schedule in which some reinforcers occurred at an IFI of 30 s and some at 240 s. The curve shows sensitivity to both periods, peaking initially at around the 30-s mark and then later around the 240-s interval. Much evidence from early research on interval reinforcement schedules suggests that animals will track probability of reinforcement as a function of time since a time marker, responding at a high rate at times when reinforcement is likely, at a lower
rate at times when it is not. Food reinforcement is perhaps the most effective time marker (cf. Staddon, 1974, 1983; Staddon & Innis, 1969), but other salient events, such as a long intertrial interval (ITI), also serve, especially if the to-be-timed interval is short (< 20 s or so, for pigeons: Starr & Staddon, 1974).

If animals are reinforced after a fixed time but also given some experience with longer times that are not reinforced, the profile of response rate vs. post-time-marker time may be highly symmetrical – especially if many trials, and data from many animals, are averaged. Figure 13.2 shows an example from an experiment with rats. The animals were trained for 60 sessions on the peak-interval procedure, a discrete-trial procedure for interval timing. The ITI was 2 s, and food was delivered for a bar press after either 10 or 20 s, signaled by a noise or a light throughout the trial. No food was delivered on half the trials; these test trials lasted for 120 s. As you can see, on the 10-s test trials, response rate peaks at about 10 s, on 20-s trials, at about 20 s. The distribution around the peak is approximately symmetrical, but response rate slowly increases towards the end of the 120-s test period. Notice also that the spread of the response-rate distributions is approximately proportional to the mean, an example of the Weber-Fechner law for the dimension of time.

The clock-like regularity of behavior on the peak procedure encouraged John Gibbon and Russell Church to make a daring proposal: that the rat is in effect a clock (Gibbon, 1977; Gibbon, Church, & Meck, 1984), with a periodic pacemaker, analogous to the pendulum of a mechanical timepiece. Within a few years, scalar expectancy theory was elaborated into an information-processing model which assumes that interval timing in animals is driven by a discrete pace-maker-accumulator (PA) mechanism that yields a linear scale for encoded time (see Gibbon, 1991, for a review).

The information-processing account raises two questions. First, how should we conceive of interval timing? As a clock-comparator system, or as a more or less analogue process whose function is to track changing reinforcement probability? The latter view is suggested by data like those in Figure 13.1. Time is a stimulus dimension like...
any other — there’s no single “clock” as such — and the important causal variable is reinforcement. This view implies that an appropriate model might be along the lines of the diffusion-generalization approach to stimulus control discussed in Chapters 11 and 12. The dimension of time is represented as a sequence of states with activation flowing either unidirectionally (Killeen & Fetterman, 1988; Machado, 1998) or bidirectionally (Staddon & Higa, 1991) from state to neighboring state. But if the clock view is correct, then the next question is: What kind of clock is it? Is it linear, as scalar expectancy theory asserts, or something else?

I am not sure whether interval timing is to be explained as stimulus control by the time dimension or as a clock-comparator. But my intuition is that the flexibility of interval timing and animals’ ability to detect more than one periodicity in the reinforcement-probability distribution argues against a single clock. However, several time-as-stimulus-dimension models have been proposed, but none is completely satisfactory. On the other hand, if there is a clock, the evidence against the assumption that time is encoded linearly seems to me strong. Despite the popularity of the SET view, I believe it to be wrong: because PA mechanisms are inconsistent with the Weber-law property of interval timing, because experiments supporting linear time can be interpreted in other ways, because SET requires too many add-on assumptions — and because there is a better alternative. In the rest of this chapter, I discuss the pros and (mostly) cons of the pacemaker-accumulator idea. The next chapter describes an alternative based on the dynamics of habituation and memory discussed in Chapter 6.

Scalar Expectancy Theory

The scalar expectancy theory of timing (Gibbon, 1977; Gibbon & Church, 1984; Treisman, 1963) was devised to explain two things: (a) that steady-state measures of time discrimination such as break point (Schneider, 1969) on fixed-interval (FI) schedules or peak-rate time (on the peak procedure) are proportional to the to-be-timed interval (proportional timing); and (b) that the standard deviation (SD) of such dependent measures is proportional to their mean (e.g., Catania, 1970; Staddon, 1965). The latter property is Weber’s law applied to the dimension of time (Weber timing). In the SET context it is termed scalar timing (Gibbon, 1977).

Scalar expectancy theory postulates a Poisson-variable “pacemaker” that begins emitting pulses a short time after the time marker. Pulses are accumulated until a short time after reinforcement, at which point the value of the accumulator is stored in a reference memory and the accumulator is reset to zero. (Partial reset is permitted, but the conditions for it to occur are not specified by SET.) Parameters are the start and stop delays and the rate of the pacemaker. Behavior is determined by comparison between the stored accumulator total and the current count. When the difference falls below a threshold (which may also vary noisily), responding, at a steady rate, begins. There is not yet consensus about the learning process: how many values are stored in reference memory, how these are selected for comparison, etc. The credit-assignment problem – how does the system “know” what stimulus to use to reset the accumulator (i.e., how does it identify the time marker? How does the system recognize a trial?) – is also left open by SET.

The assumption that operant responding on interval schedules is essentially an on-off affair – the animal either responds at a steady rate or not at all – “break-and-run” – is based on data from Schneider (1969) who analyzed the “scallops” generated by well-trained pigeons on a range of FI schedules. Schneider used an algorithm to identify the point in each interfood interval where response rate changed from “low” to high.” Data from all intervals were then superimposed at this point, yielding an average that showed an apparently sharp transition from low to high. Gibbon (1991) concludes that “since the ‘break point’ [i.e., the point of transition from low- to high-rate responding] varied from trial to trial, it generated the smooth, scalloped pattern when averaged.” (p. 11) But the implication that responding on individual trials is break-and-run is false, as the real-time cumulative record in Figure 2.7 clearly demonstrates. (Indeed, the term
“scallop” was first used in the 1950s as a description of individual cumulative records, not averages.) The curvature is perfectly real, at least at longer FI values and early in training at shorter values.

Schneider’s own analysis provides only qualified support for the two-state idea. Figure 13.3 shows Gibbon’s replot of Schneider’s break-point data, and it is evident to the eye that at the two longest FI values, responding slowly accelerates from zero to a maximum value (arrows). These results are perfectly compatible with the assumption that response probability increases smoothly, rather than changing abruptly, across the to-be-timed interval\(^ {107} \) (Box 13.1).

Timing with a pacemaker/accumulator implies greater relative accuracy at longer time intervals. If there is no error in the accumulator, or if the error is independent of accumulator value, and if there is pulse-by-pulse variability in the pacemaker rate, then relative error (SD divided by mean, coefficient of variation: CoV) must be less at longer time intervals. The usual theoretical relation is that the SD is proportional to the square root of the mean.

This relative improvement with absolute duration does not depend on the type of variability in the pacemaker, so long as it is independent of time. But in fact, CoV is approximately constant (Weber’s law, the scalar property) over a limited range of times. At longer times the CoV tends to increase (rather than decrease) with the duration of the timed interval (e.g., Gibbon, Malapani, Dale & Gallistel, 1997; Zeiler, 1991). This problem was apparent early on in the evolution of SET, but the pacemaker-accumulator feature was nevertheless retained.

But it isn’t used. For example, Gibbon suggests that the Poisson pacemaker has a “high rate” (1991, p. 22) so that “Poisson [pacemaker] variance rapidly becomes swamped by scalar variance [i.e., a noisy multiplier] (Leak & Gibbon, 1995, p. 18).” Gibbon (1992), has shown how assumptions about memory encoding and decoding “allow multiplicative random variables to intervene…between the value of a count in the accumulator at reinforcement time and its value after retrieval when it is used for comparison (p. 289).” Thus, the increasing relative accuracy at longer times intrinsic to any pacemaker-accumulator system is irrelevant, because pacemaker rate variation makes a negligible contribution to total variance. In current versions of SET, the Poisson pacemaker assumption is redundant: either the Poisson property is replaced by a constant rate that varies from trial to trial; or, alternatively, residual Poisson variability is deemed to make a negligible contribution to total variance. In a recent commentary, Gibbon concedes that “The key features of the theory did not rely on the Poisson pacemaker idea…” (Gibbon, 1999, p. 272)

\(^ {107} \) Machado (1997) has proposed a real-time model for FI acquisition that accommodates the “scallop” data.
Box 13.1: Interval Length and the FI “Scallop”

The apparent failure to find a scallop – gradual acceleration of response rate through the to-be-timed interval – at small FI values may be an artifact. The break-and-run pattern at short intervals may be a consequence of an apparently unrelated fact: that in well-trained animals, overall response rate (responses per minute over the whole interval) is more or less constant over a range of intervals, i.e., much the same at FI 15 s as FI 120 s. An approximately constant response rate means that there will be fewer response opportunities in short intervals than in long ones. Limited sampling of the underlying process at short intervals may account for the break-and-run pattern.

The argument is as follows. Let’s assume that response probability in fact increases gradually, rather than abruptly, across the to-be-timed interval:

\[ p(x) = \frac{kt}{I}, \quad B13.1 \]

where \( p(x) \) is the probability of a response in any time step, \( t \) is post-food time and \( I \) is the interval duration. The equation scales each interval duration so that response probability is zero at the beginning of the interval and one at the end.

Suppose that the occurrence of a response depends on two factors: response opportunity and response probability:

\[ \text{If } p(x) \times t/I, \ x = 1, \text{ otherwise } x = 0, \quad B13.2 \]

where \( x = 1 \) means respond and \( x = 0 \) means no response, at a given response opportunity. And assume that response opportunities occur at a constant rate, so that there will be fewer response opportunities in a short interval than in a long. Consequently, \( p(x) \) is sampled less often in a short interval than in a long. In other words, suppose that response opportunities come along at fixed intervals of time and during each response opportunity, response probability is computed, according to Equation B13.1. If \( p(x) \times t/I \) (Equation B13.2) a response occurs during that response opportunity; otherwise, there is no response.

What kind of cumulative records will be generated by such a system? Will they be smooth (mirroring the smooth increase in response probability across the interval); or will they be break-and-run (mirroring Schneider’s data at short intervals)? Figure B13.1 shows two typical simulation results. The conclusion is that when there are few response opportunities in an interval, the typical cumulative record will show little acceleration. It will approximate break-and-run (light line). But when there are many response opportunities, the record becomes more scalloped (heavy line). There are certainly other continuous models that will generate this result. Conclusion: the data – scalloped cumulative records at long intervals, break-and-run at short – are perfectly compatible with an underlying process in which response probability increases smoothly across the to-be-timed interval.

The core assumptions of SET are the following. SET, and perhaps any “clock” theory of operant timing, requires three time-related variables: real elapsed time, the encoded value of current time and the remembered value for times encoded in the past. We denote real time as \( t_i \), where \( i \) indicates the relevant time marker. The encoded value for \( t_i \) is \( \tau_i \). The remembered value for a past \( \tau_i \) I indicate by bold italic: \( \tau_i \). The superscript * denotes the value of each variable that is associated with reinforcement. The internal variable for current time, \( \tau_i \), is always referred to in SET as “number of pulses,” but because the PA assumption is in fact unnecessary, it could simply be any variable proportional to real time.

SET assumes that the relation between remembered time and real time is linear. Formally, \( \tau = kt \), (subscripts neglected for simplicity); that is, remembered time, \( \tau_i \) is proportional to real time, \( t_i \), and \( k \) is a constant of proportionality. But this cannot be a direct relation, because remembered time, \( \tau_i \) is not stored directly – what is stored is encoded time, \( \tau \). So, the correct relation is:

\[ \tau = k\lambda t, \quad (13.1) \]

175
where $\lambda$ denotes the pacemaker rate — in effect, the scale relation between (linearly) coded internal time $\tau$ and real time $t$. Variation in pacemaker rate, $\lambda$, is assumed to be negligible; hence the major contribution to these distributions is variation in parameter $k$. Since $k$ is a multiplier, any variation in $\tau$ will automatically be proportional to $t$, i.e., scalar, in SET terminology. Thus, SET explains the scalar property by assuming (a) that there is a multiplicative transformation between encoded time and remembered time; (b) that temporal judgements represent a comparison between (long-term) remembered time and (short-term-encoded) current time; and (c) that most of the variability in remembered time is due to the multiplicative relation ($k$) between encoded and remembered time. The Poisson pacemaker-accumulator system is completely redundant.

**Response Rule.** Predictions of response pattern vs. time (and sources of variability) are obtained through a threshold assumption:

$$\text{if } |\tau^* - \tau| < \theta, \text{ response rate } = x$$
$$\text{otherwise, response rate } = 0,$$

(13.2a)

where $x$ is a constant and $\theta$ a threshold. (The constant-rate assumption is not accurate, as I just pointed out, but it allows the clock analysis to go forward.) Because both $\tau$ and $\tau$ are linear with respect to real time (Equation 13.1), $t$ may be substituted so that Equation 2a is thus shorthand for ($\lambda$ cancels)

$$\text{if } |kt^* - t| < \theta, \text{ response rate } = x$$
$$\text{otherwise, response rate } = 0.$$

(13.2b)

In practice, however, a ratio version of Equation 13.2b is preferred:

$$\text{if } |(kt^* - t)/kt^*| < \theta, \text{ response rate } = x$$
$$\text{otherwise, response rate } = 0.$$

(13.2c)

Equation 13.2c, linear time encoding with a ratio response rule, is equivalent to logarithmic time with a difference response rule:

$$\frac{(kt^* - t)}{kt^*} < \theta = t/kt^* > 1 - \theta = \ln t - \ln t^* > \theta,'$$

(13.2d)

where $\theta'$ is a threshold value. Nevertheless, the ratio-rule-with-linear-time version is preferred in SET because of the commitment to a pacemaker.

The essential features of SET are thus relatively simple: linear encoded and remembered time, related multiplicatively, and all-or-none behavior generated via a thresholded comparison between them. In Chapter 14, I will argue that the data better support approximately logarithmic time encoding, as Equation 13.2d implies.

**Alternatives to Linear Encoded Time**

The only necessary requirement for an interval-time clock is some internal variable that changes in a reliable monotonic way with time elapsed since a time-marker. Moreover, so long as there is a unique value of the variable for each time, it makes no difference whether the variable increases or decreases with time. Given a monotonic function, time can be told by associating specific values of the variable with reinforcement or its absence and responding accordingly. All “clock” theories of interval timing compare some internal time-related variable, $\tau$ =

---

108 So long as the variation is between trials not within trials (see Staddon & Higa, 1999, for a full account).
109 Or even a covert external variable: [https://www.youtube.com/watch?v=Ftr9vY-YuYU](https://www.youtube.com/watch?v=Ftr9vY-YuYU)
Adaptive Dynamics

Staddon

\[ f(t), \text{ with the (remembered) value of that variable at the time of reinforcement. There is no reason to assume a priori that } f(t) \text{ is linear. An obvious nonlinear possibility is the logarithmic function:} \]

\[ \tau = k_1 - k_2 \ln(t), \quad (13.3) \]

110 Since it doesn’t matter whether \( f(t) \) is increasing or decreasing, I use the decreasing form of log function because I later introduce the idea of a decreasing memory trace as a basis for \( f(t) \).

where \( a \) and \( k_i \) are constants. Gustav Fechner showed many years ago that a logarithmic function is consistent with Weber’s law. If internal noise, \( \varepsilon \), is independent of \( \tau \) (i.e., \( \tau = k_1 - k_2 \ln(t) + \varepsilon \)), then Weber’s law (the scalar property) can be obtained directly, just from the form of \( \tau \), because the slope of \( \text{Equation 13.3} \) is inversely related to \( t \):

\[ \frac{d\tau}{dt} = -\frac{k_2}{t}. \quad (\text{Given a constant variation in response threshold, therefore, variation in the time of onset of responding (Equation 13.2) will be proportional to the slope of } f(t), \text{ hence (for the log function) proportional to } t. \]

Temporal Bisection. The log-time assumption is consistent with temporal bisection data, which show that animals judge an event of duration \( x \) to be equally like two comparison durations \( y \) and \( z \) if \( x \cong \sqrt{yz} \), i.e., at the geometric mean. In a typical bisection experiment (e.g., Staddon, 1984; see Figure 13.4) the organism has two choices and is presented on each trial with one of two stimulus durations, \( T_S \) and \( T_L \). Reinforcement is delivered for Response A following the short duration, \( T_S \), and Response B following the long, \( T_L \). In occasional probe trials, intermediate durations are presented. The typical finding is that subjects are indifferent between A and B when the probe duration, \( T_P \), is equal to the geometric mean of \( T_S \) and \( T_L \): \( T_P = (T_S.T_L)^{1/2} \). This is what would be expected given symmetrical variation around \( T_S \) and \( T_L \) on a logarithmic psychological time scale: Responses A and B should have equal strength at the geometric mean of the short and long training times. This result is different from what would be predicted by SET. For example, given \( T_S = 1 \) and \( T_L = 2 \), logarithmic time predicts bisection at \( T_P = 1.414 \) and scalar timing at \( T_P = 1.33 \), the harmonic mean (linear time would place the point of subjective equality [PSE] at 1.5).

There are several nonlinear functions that have very similar properties to the logarithmic: power (e.g., Staddon, 1984), the sum of exponentials (multiple time scale: MTS\(^{111}\)) and others. The logarithmic, power and MTS functions can approximate both Weber’s law and necessary memory properties such as Jost’s law. The power function is the candidate offered by Wixted and Ebbesen (1991, 1997) for the forgetting function in a range of species. The log and power functions were the best fits to the large human-forgetting-curve dataset reviewed by Rubin and Wenzel, 1996 (they did not look at MTS-type functions). The resemblance between the MTS

111 But note that because of the nonlinear dynamics of the MTS model, the form of the function is not invariant, but depends on the system history – see Chapter 14.
and power functions will become relevant when I present a memory-based approach to time discrimination in Chapter 14.

**Argument against Nonlinear Time**

Given the apparent equivalence between models with a linear time scale and a ratio-difference response rule vs. a logarithmic scale and a constant-difference response rule, it is obviously important for linear-time models like SET to rule out the hypothesis that time is encoded nonlinearly.

**The Time-Left Experiment.** The main argument against nonlinear encoding is based on “time-left” choice experiments. In the time-left procedure, rats or pigeons must choose between a short fixed delay and a longer fixed delay that has already partly elapsed (“time left”; see Figure 13.5). In Gibbon and Church’s (1981) Experiment 1, for example, rats chose between two levers, one fixed, the other retractable, representing different delays until food reinforcement. The fixed lever signaled a fixed delay, C, (timed from trial onset) until food. The delay on the retractable lever, S (timed from lever presentation), presented T s into the trial, was also fixed, but shorter than C. The experimental question is: how will preference for fixed delay S vs. time-left delay C-T change as a function of T? In particular, will the animals prefer the time-left lever when C-T=S? In fact, rats are indifferent when delays to food are equal on both levers, which Gibbon and Church take as evidence against logarithmic coding of time: “A logarithmic or indeed any curvilinear subjective time scale ought to result in a greater preference for the time-left side of the choice when the real time left to food is equal on both alternatives.” (p. 92).

The argument against this conclusion has been presented at length elsewhere (Staddon & Higa, 1999; Staddon, Higa, & Chelaru, 1999) and I just summarize it here. The flaw boils down to a confusion between subjective time and encoded time. Subjective time is both *qualia,* how time feels, and objective judgement. About qualia, science has nothing to say. But we can ask a man, or a pigeon, whether the first 60 s of a 120-s interval appears to be more or less than the second 60 s. The answer is clear: the first and second 60 s will be judged more or less equal by both man and beast.

Encoded time is measured differently. Here the issue is how well can one time interval be discriminated from another. The assumption is that the encoding process will impose discriminability limits, so that knowing the limits, we can infer the process. Here also, the answer is clear: Weber’s law holds approximately over a wide range, which implies nonlinear encoding of time that is approximately logarithmic. Logarithmic temporal encoding does not conflict with linear temporal judgment for the same reason that accurate judgments of spatial distance do not conflict with foreshortening on the retinal image. Distance is encoded nonlinearly on the retina: distant objects have smaller retinal images than nearby ones. But downstream processes decode the retinal image in a way that allows distance judgments to be more or less veridical. A meter rod looks to be a meter long whether it is 10 cm from us or 100. At the same time, we are better able to discriminate two objects a meter apart when they are 3 meters away than when they are a
100 meters away. In the language of psychophysics, confusion scales are not the same as magnitude-estimation scales (Stevens, 1975).

**Argument for Nonlinear Time.**

The time-left experiments are not a convincing argument against nonlinear encoding of time. Is there any evidence in favor of log-like encoding, beyond the Weber-law standard-deviation property and geometric-mean bisection data? There is some indirect supporting evidence that derives from the widespread finding of power-law relations in interval-timing experiments. Platt (1979) reviewed numerous studies showing a power-law relation between temporal dependent and independent variables in temporal differentiation and discrimination experiments:

\[ b = qt^s, \]  

(13.4)

where \( b \) is the observed behavior (e.g., response duration, waiting time), \( t \) is the required duration, \( q \) is a constant and \( s \) is an exponent (usually close to one). Power function relations with exponents different from unity cannot easily be reconciled with SET, but there is a theoretical argument that ties them to logarithmic internal coding. The argument is as follows.

First assume that temporal discrimination is a comparison process in which an internal, logarithmic temporal variable (reference memory) is compared with an output variable (working memory: encoded elapsed time) that is also logarithmic. In other words, if memory for a time interval is encoded logarithmically and if current elapsed time is also encoded logarithmically, and if behavior involves responding when the two are equal or differ by a constant, then the empirical relation between temporal independent and dependent variables will take the power form.

This interpretation of the psychophysical power law by Ekman (1964) and MacKay (1963) was shown to fit psychophysical data by Staddon (1978) and the argument can be applied to time discrimination. I assume that the internal effects, \( dz \), of both remembered time (represented in Equation 13.5) and elapsed time (represented by Equation 13.6) show Weber-law sensitivity, according to sensitivity coefficients (Weber fractions), \( w_t \) and \( w_b \):

\[ \frac{dt}{t} = w_t dz_t \]  

(13.5)

\[ \frac{db}{b} = w_b dz_b . \]  

(13.6)

The first equation simply states that a small change in real time, \( dt \), has a psychological effect, \( dz \), that is inversely related to \( t \) and the Weber fraction \( w_t \): \( dz = dt/w_t t \); and similarly for the second equation (Staddon, 1978).

Integrating both sides of Equations 13.5 and 13.6 yields

\[ \ln t + K_1 = w_t z_t \]  

(13.7)

\[ \ln b + K_2 = w_b z_b, \]  

(13.8)

a logarithmic relation between both remembered time, \( t \), and elapsed time, \( b \), and their internal effects, \( z_t \) and \( z_b \). \( K_1 \) and \( K_2 \) are constants of integration. In temporal-discrimination experiments, the internal effects of remembered and elapsed time are equated, \( z_t = z_b \), which allows us to eliminate \( z \) from Equations 13.7 and 13.8. Rearranging yields the power relation (Equation 13.4), with

\[ q = \exp(w_b K_1/w_t - K_2) \]  

(13.9)

and

\[ s = w_b/w_t. \]  

(13.10)
The two constants, $K_1$ and $K_2$, are scale factors, assumed to be constant across different experimental procedures.

Notice that if the sensitivities (Weber fractions) of remembered time and elapsed time are the same, the exponent, $s$, is unity and behavior (waiting time), $b$, is linearly related to elapsed time, $t$. This is a common, but not universal, result in temporal experiments. The exponent for the function relating waiting time to fixed-interval duration in steady-state parametric FI experiments is usually close to one. But the exponent in steady-state tracking experiments, in which the animal is repeatedly subjected to cyclically varying interfood intervals, is typically less than one. This is just what we would expect, given that the exponent $s = w_t/w_b$ and that it is harder for the animal to remember the upcoming interfood interval when several are intermixed in each session than when all the intervals are the same from session to session. If $w_t$, the Weber fraction for remembered time, increases (i.e., poorer discrimination) then the exponent $s$ should decrease.

As this argument suggests, Innis and Staddon (1971) found a less-than-one power-function exponent of .824 in an early interval-tracking experiment in which pigeons were repeatedly exposed to a cycle of 7 ascending and 7 descending interfood intervals. They also found that the exponent increased to .894 when different discriminative stimuli signaled the ascending and descending parts of the cycle and presumably reduced memory interference among remembered intervals.

If different experimental arrangements affect only sensitivities and the two sensitivities are affected differentially, then the power function exponent will be different in different experiments. It follows from Equations 13.9 and 13.10 that the slopes and intercepts of a set of such functions will be linearly related

$$\ln q = sK_1 - K_2,$$

which is a testable empirical prediction.

DeCasper (1974, cited in Platt, 1979) plotted the slopes and intercepts of power functions obtained in four different temporal differentiation experiments, and the slopes and exponents show a reasonably good linear fit to Equation 13.11 (cf. Staddon & Higa, 1999, Figure 2). Taken all together, the data and arguments in favor of time coding that is approximately logarithmic are stronger than for any other simple function.

**Individual Trials: Correlation Data and Distributions**

In the SET analysis of individual trials on the peak procedure (Church, Meck, & Gibbon, 1994), well-trained rats are said to learn on each trial to wait for a certain time (the *start time*) until they begin pressing the lever at a constant rate, which they maintain until they eventually stop responding (at the *stop time*). I have already pointed out that the constant-rate assumption is only approximately true. In addition to the plain fact of scalloped FI cumulative records, Cheng and Westwood (1993) present peak-procedure data showing that “within the run phase was an inner run phase at a yet higher rate” (p. 56). And much earlier, Staddon and Frank (1975) showed that on fixed-
interval schedules, pigeons respond initially faster, and accelerate faster, the later they begin responding within an interval (Figure 13.6). Timing data have a complex structure.

Nevertheless, if we accept for the moment the stop-run-stop assumption and the threshold assumption (Equation 13.2, above) – together with the scalar assumption and the assumption of linear encoded time – then SET makes some clear predictions about the correlations between start time ($t_{\text{start}}$), stop time ($t_{\text{stop}}$) and spread ($t_{\text{stop}} - t_{\text{start}}$) on the peak procedure. These predictions are illustrated in Figure 13.7. The figure shows the value of the difference between $\tau$ (encoded time) and $\tau^*$ the remembered time at which reinforcement occurs, (corresponding to $T$ in real time, when there is no reference-memory variance) as $t$ increases from zero, through $T$ (the time of reinforcement) and beyond (V-shaped pair of lines). The threshold is assumed by SET to vary, and two threshold values (horizontal dotted lines) are shown, with two start and stop times. When the threshold is low, start time is long and the time between start and stop is short; conversely, when the threshold is high, start time is short and the time between start and stop times is long. In SET, therefore, the correlation between $t_{\text{start}}$ and both $t_{\text{stop}}$ and $t_{\text{stop}} - t_{\text{start}}$ (spread) is negative.

The data agree only partially: the correlation between start and spread is indeed negative, as SET, and some other theories (Killeen & Fetterman, 1993; Staddon, 1972a; Staddon and Ayres, 1975) predict (mean: -0.33, Church, Meck & Gibbon, 1994, Table 2). But the correlation between start and stop is in fact positive (mean: 0.31), as predicted by response-competition theories (e.g., Staddon & Ayres, 1975), not negative as predicted by this version of SET. Church et al. (following Gibbon and Church, 1990) interpret the positive start-stop correlation as evidence for memory variance (the once-per-trial sampling of $\tau^*$, discussed earlier) plus different stop and start thresholds, i.e., they reconcile model and data by adding another process and parameter. In their final statement on this problem, Church, Meck & Gibbon (1994) conclude that the covariance data support a “scalar timing model in which animals used on each trial a single sample from memory of the time of reinforcement and separate response threshold to decide when to start and stop responding. (p. 135)” – a total of five parameters. Even without the second threshold Gibbon & Church (1990) concluded in their earlier paper “The flexibility of the surviving [SET] models is of course bought at some cost, since additional assumptions ineluctably lead to additional parameters.” (p. 53). Perhaps for this reason, the single-threshold, linear-difference, no-memory-variance model continues to be used (e.g., Leak & Gibbon, 1995, Fig. 1) despite its inconsistency with the correlation data.

Figure 13.7 makes other predictions about peak-procedure data. Most discussion is devoted to peak-rate distributions (e.g., Figure 13.2) which are generally Gaussian, symmetrical, scalar and centered on the time of reinforcement (but see Cheng & Westwood, 1993; Zeiler & Powell, 1994). The picture for start and stop times is not so simple, however. For example, if the only source of variance in the model is variation in the single threshold, then, because of the linear time assumption, distributions of start and stop times should be symmetrical with the same
standard deviation. They are neither. Church et al. report that the SD of start times is less than the SD of stop times (for the same T value: see their Figure 8). At least one other study seems to show larger SD for start than for stop times, and the start distribution is highly skewed: Brunner et al. (1997, Figures 2, 8 & 11) report start-time data that are strongly skewed to the right, not symmetrical. The start-time data in Figure 13.6 (inset) are also skewed to the right.

Admittedly, the definition of “start” in the peak procedure, and on fixed interval, is somewhat ambiguous, but on response-initiated-delay (RID) schedules it is not. On RID schedules, the first response after reinforcement causes a stimulus change and initiates a delay, T, until the next reinforcement, which requires no additional response. On so-called clamped schedules, the time between reinforcement is fixed, so that t+T is constant, where t is the wait (start) time until the first response (i.e., the procedure resembles an FI schedule). Start-time distributions on such procedures are symmetrical on a log (not linear) scale (see Figure 13.8). This result is consistent with the MTS approach (discussed in Chapter 14), but not with SET.

Zeiler and Powell (1994) report that when only individual-subject data are considered, start and stop times (measured by Schneider’s [1969] break-point procedure) show constant coefficient of variation, consistent with the scalar property, but pause (time to first response — start time) and peak-rate time (the preferred variable in SET peak-procedure experiments) show CoV increasing as a function of schedule fixed-interval value. They also show a number of power-function (rather than linear) relations between temporal dependent and independent variables (as have a number of other investigators, cf. Platt, 1979). Zeiler and Powell conclude “Existing theory does not explain why Weber’s law so rarely fit the results or why each type of behavior seemed unique.” (p. 1)

Summary

This discussion of SET illustrates two quite different approaches to theorizing. One approach, common in information-processing models and among “cognitive” theorists generally, is to assume, on a priori or intuitive grounds, the existence of some mental process — such as a pacemaker-driven internal clock. Research then proceeds in two steps. The first step is to refine an experimental paradigm that is presumed to provide a “pure” measure of clock properties (the peak procedure, with suitable choices of probe-trial probability and duration, and intertrial-interval distribution). The next step is to derive exact predictions about the quantitative properties of behavior in this situation. If these are confirmed, the theory is supported. If not, the theory is modified or new properties of the clock are assumed to have been revealed. For example, the theory may be modified by adding nonlinear elements (thresholds, multipliers) or sources of variance, until the new data are accommodated. Or, if timing falls short, it is assumed that only partial or delayed reset has occurred: “[this result] suggests that an animal can reset its internal
clock in less than 2 sec...This suggests that resetting of the internal clock occurs in a non-time dimension.” etc. (Church, 1980, p. 208) The core assumption – that there is a clock, quite separate from other processes that underlie learned behavior – is never directly tested.

Cognitive theory frequently shows little concern for uniqueness. For example, the SET literature abounds with phrases like “Here, we describe the version of the model used in fitting the data from the two experiments.” (Allan & Gibbon, 1991, p. 41) “I consider here two response rules, both of which have been used in different contexts…” (Gibbon, 1991, p. 25) or “I begin with a general model which applies to any of the timing performances…but with differing instantiations in the components.” (Gibbon, 1991, p. 21) The model is assumed to be true, and the task of experiment is identify sources of variance within the model framework: “The estimate of a time interval may reflect sources of variance in the clock, memory, or comparator.” There may be some acknowledgement that “The estimate may also reflect sources of variance that are independent of the timing process (Church, Miller, Meck & Gibbon, 1991, p. 207).” – but the assumption that there is a separate “timing process” is never questioned. The model itself is treated like a set of interchangeable parts that in different combinations can fit data from almost any interval-timing experiment. The usual scientific strategy of looking for the theory for a given set of phenomena has been abandoned. Instead, the object is some explanation for each phenomenon – within the flexible range of possibilities defined by SET.

This is not necessarily a bad strategy – after all, living organisms are also constructed from a limited number of parts assembled in different combinations. The SET approach has also been helpful in designing and organizing experiments and can give a sense of closure to the analysis of puzzling data. But it has pitfalls. First, the initial PA-clock hypothesis was invented entirely a priori. No one had seen an interval pacemaker, and none has since been demonstrated even after twenty years of search. The existing data do not demand a pacemaker. The idea is attractive because it is familiar, and because it is analyzable, not because it explains anything in the sense that you get more out of it than you put in. Second, because the approach is so flexible and the data set is relatively limited, the research program does not – and perhaps can not – lead to abandonment of the original theory, which can almost always be preserved via modification and elaboration.

What remains of SET, once the pacemaker has been abandoned, is the idea that interval timing involves some kind of comparison between (encoded) elapsed time and remembered time. I have argued that linear encoding is unlikely. What is the alternative? Chapter 14 takes up this theme.
Chapter 14

TIME, II

How noiseless falls the foot of time! (W. R. Spencer: Lines to Lady Hamilton.)

The proximal causes of behavior in interval-timing experiments are complex. Measures such as breakpoint and wait time are not determined just by post-time-marker time. For example, experiments have shown proactive interference — effects on wait time of events preceding the time marker (Staddon, 1974, 1975). Other experiments have shown apparently cumulative effects of successive time markers (Hornr, Staddon & Lozano, 1997). Activities other than the operant response may need to be taken into account. On periodic food-reinforcement schedules, rats, pigeons and many other animals engage in a variety of interim activities in addition to the terminal, operant response (Lejeune, Cornet, Ferreira, & Wearden, 1998; Machado & Keen, 1999) Shettleworth, 1975; Staddon & Simmelhag, 1971; Timberlake, 1983). These activities seem to compete, with each other and with the operant response (Staddon, 1972a, 1977; Staddon & Frank, 1975). For example, restrained pigeons wait a shorter time until the first

<table>
<thead>
<tr>
<th>Rat</th>
<th>A-B vs. B-C</th>
<th>A-D vs. D-E</th>
</tr>
</thead>
<tbody>
<tr>
<td>RH 2</td>
<td>-.34*</td>
<td>-.44**</td>
</tr>
<tr>
<td>RH 3</td>
<td>-.30</td>
<td>-.87</td>
</tr>
<tr>
<td>RH 4</td>
<td>.11</td>
<td>-.32**</td>
</tr>
<tr>
<td>RH 5</td>
<td>.18</td>
<td>-.42**</td>
</tr>
<tr>
<td>RH 6</td>
<td>-.49**</td>
<td>-.66**</td>
</tr>
</tbody>
</table>

Table 14.1 Correlations between start time and duration of drinking (AB vs. BC) and between start time and duration of wheel running (AD vs. DC) for five rats after long exposure to periodic food delivery in the hexagonal apparatus shown in Figure 14.1. *: p < .05; **p < .01. (from Staddon & Ayres, 1975, Table 4).

Figure 14.1 Top: Hexagonal apparatus used by Staddon & Ayres (1975) to study time-allocation by rats on periodic-food schedules. Bottom: Points of transition between successive activities whose correlations are given in Table 14.1.
postfood fixed-interval response than unrestrained animals that can move around and engage in interim activities (Frank & Staddon, 1974). An interim activity like wheel running seems to delay the terminal response of lever pressing – omitting the wheel causes lever pressing to begin earlier in the timed interval (Staddon & Ayres, 1975).

Each activity has a degree of autonomy and each may also be sensitive to temporal cues. In a well-trained animal, each activity seems to occur for a characteristic time and the time between the end of one activity and the beginning of the next is shorter the later the ending time of the first activity (cf. Staddon, 1972; Staddon & Ayres, 1975, Table 4 — see Figure 14.1). Consequently, when an activity begins late in the interfood interval it tends also to end late. This may be the explanation for the positive start-stop correlations in peak-procedure experiments (Church, Meck & Gibbon, 1994). Conversely, when an activity begins early, it can continue for longer, because the next activity in the sequence (which has its own temporal causal factors) doesn’t cut in until its proper time (Figure 14.1). This may account for the negative start-spread correlations discussed earlier. In sum, even on strictly periodic reinforcement schedules, postfood time is only one of several causal factors acting on the operant response.

Most recent experiments on interval timing have not favored strictly periodic schedules, however. The modal study is more likely to use something like the peak procedure (perhaps with two or three response alternatives), or some kind of comparison method such as bisection, rather than a simple fixed-interval schedule. The aim is to assess the “time sense” in a way that parallels human psychophysical measurement. But under these conditions the identity of the time marker is often uncertain. As Stubbs, Dreyfus, Fetterman, Boynton, Locklin and Smith (1994) point out, “with complex stimulus arrangements, like those used in much current nonhuman animal research, multiple aspects of complex stimuli affect behavior and complex stimuli exert multiple effects on behavior.” (P. 31) Given the complex response and stimulus properties of interval-timing experiments, it may be a mistake, therefore, to expect the kind of exclusive control by time that is taken for granted by cognitive timing theories. Consequently, it may be unrealistic to insist on quantitative precision from a theory that deals only with a single activity, such as pecking or lever pressing, and only with time. Bearing these limitations in mind, let’s look at a memory-based approach to temporal control (Staddon & Higa, 1999).

MEMORY TRACE AS A TIMER

As I pointed out in the previous chapter, it is by no means certain that interval timing involves any kind of direct pair-comparison between (encoded) current and remembered time. Nevertheless, this model may fit well enough what is happening on strictly periodic schedules (e.g., FI), the main topic for the following discussion. Scalar expectancy theory assumes that the first step in such a comparison, the encoding, is linear: encoded time is proportional to real time. But as I argued earlier, the SET case for a linear clock is not conclusive. In every other sensory dimension that obeys Weber’s law, the encoding is compressive: a large physical range is compressed into a much smaller physiological range, and Weber’s law — error is proportional to the mean — is the result. There is a strong prima facie case, therefore, for a compressive, hence nonlinear, form for \( f(t) \), the function relating real and encoded time. Possibilities for \( f(t) \) are the power and MTS functions, that share with the log function the property that they decline rapidly at first and more slowly later (e.g., Dreyfus, Fetterman, Smith & Stubbs, 1988; Staddon, 1983, 1984 & 1997a; Stubbs, et al., 1994). Which function is best?

The MTS function I discuss in this chapter (see also Chapter 6) is derived from data on habituation and memory. It is intrinsically dynamic. It implies highly compressive, logarithmic-like temporal encoding; and, unlike SET and other static timing theories (cf. Machado, 1998), it suggests ways to deal with assignment of credit. It implies that there may be no separate “internal clock” at all; that interval time discrimination is just like any other discrimination. The main difference is what is discriminated. In time discrimination, I will argue, animals are learning to
discriminate between memories of different ages and thus of different “strengths.” This approach to timing offers the possibility of integrating a wide range of phenomena – habituation, rate-sensitivity, partial-reinforcement effects, Jost’s law, Weber-law timing, the psychophysics of temporal choice and many properties of time-based reinforcement schedules (cf. Staddon, 1993, 1997a; Staddon & Higa, 1996, 1999) – by means of a single event-memory process.

The Role of the Time Marker

All interval timing is timing with respect to some stimulus, such as food reinforcement or the intertrial interval, that acts as a time marker. (Multiple stimuli may act as time markers under some conditions, but I don’t deal with that possibility here.) The time marker initiates the interval that is to be timed. The time marker, and the different properties of different kinds of time marker, is largely ignored by cognitive timing theories, even though there is abundant evidence that timing is more effective with some time markers than others. Reinforcement, or a stimulus that signals reinforcement, is more effective than neutral stimuli, for example (Roberts & Holder, 1984; Spetch & Wilkie, 1981; Staddon & Innis, 1969; Starr & Staddon, 1974).

A time marker, like any stimulus, affects the state of the organism. Some effects may be more or less permanent, but other effects continue to change with time, typically to grow weaker as memory for the event “decays.” (There are also interactions among memories — retroactive and proactive interference — but I focus here on the decay aspect.) If the organism is sensitive to the changing aspects of a memory; if we can discover a quantitative form for the change in the memory variable with time; and if the organism can learn to associate specific actions with specific values of the memory variable, then we have provided the organism with a potential “interval clock.” Moreover, we have done so by making use of a familiar process – event memory112 – rather than by postulating an ad hoc pacemaker-accumulator. It is always possible that there is an internal clock independent of memory processes. But it is more parsimonious to see first how well we can explain interval timing with known processes, before resorting to an assumption whose main basis is the phenomenon it is supposed to explain.

So, what is an appropriate experimental paradigm for event memory? And, what does this paradigm tell us about the way that stimulus effects change with time?

The Event-Memory Paradigm.

The most elementary effect of stimulus presentation is habituation, the waning of a reflex response to successive presentations of a stimulus. I assume that stimuli that elicit no reflex response are nevertheless remembered in much the same way as stimuli that do elicit a response. The evidence includes phenomena such as latent inhibition, that demonstrate memorial effects of “neutral” stimuli. Hence, the habituation experiment is the most elementary manifestation of event memory and the dynamics of habituation can tell us how the effects of a stimulus change with time.

As we saw in Chapter 6, habituation is widely observed with many different responses and stimuli and across species ranging from protists to humans. Rate sensitivity is the surprising fact that although habituation is more rapid and complete when interstimulus intervals (ISIs) are short than when they are long, recovery is also more rapid after short ISIs. Rate sensitivity puts helpful constraints on the dynamics of event memory.

Chapter 6 described how habituation can be duplicated by a process in which response strength is the difference between a constant stimulus effect and a leaky-integrator short-term

112 I prefer the term event memory, rather than the more familiar short-term memory (STM), because the term need not in fact be short — event memory can extend over weeks. Since the distinctions between the various types of memory are still not well understood, the name is not critical providing we understand what is meant in a particular context.
stimulus memory (Figures 6.4 & 6.6). As successive stimuli are presented, the accumulated charge of the integrator – event memory for the stimulus – increases. Since response output is the difference between the integrator charge and the constant stimulus input, output response strength decreases as successive stimuli are presented and memory strength builds up. If the ISI is too large, however, the integrator “discharges” (“forgets”) in between stimulus presentations and the system does not habituate. Thus, the degree of habituation in the model, as in nature, is less at longer ISIs.

The single-unit model habituates less at long ISIs than at short, but post-habituation recovery is just as rapid – because only one time constant is involved. As I pointed out in Chapter 6, a series of at least two cascaded habituation units (the output of the first being the input to the second, and the second slower than the first) is necessary to reproduce the rate-sensitive property (earlier habituation models by Gingrich and Byrne, 1985, and Treisman, 1984, also incorporate two stages). Staddon and Higa (1996) showed how this type of model can be extended to any number of stages.

Memory Trace Strength. Spontaneous recovery, the increase in reflex strength as a function of time in the absence of further stimulation, is the reflex counterpart of forgetting. For a habituated reflex, response strength recovers with time as memory for the stimulus fades. For a memory, strength decreases with time. In the cascaded-integrator model, memory trace strength – recallability of a stimulus – is thus the complement of reflex strength:

\[ \mathcal{V}_M = 1 - \mathcal{V}_R = \sum_{j=1}^{N} V_j, \]

where \( \mathcal{V}_R \) is reflex strength, \( \mathcal{V}_M \) is memory trace strength, \( V_j \) is the activation level of the \( j \)th integrator, the stimulus strength is unity and \( N \) is the number of cascaded habituation units. A realistic model for interval timing seems to require about ten units.

**MTS-Memory Timing Model**

The basic idea of the multiple-time-scale (MTS) model is that what is learned on periodic schedules is the reinforced and nonreinforced values of the animal’s memory trace for the time marker, i.e., the values of memory trace strength, \( \mathcal{V}_M \), that are associated with reinforcement and nonreinforcement\(^{113}\). One objection to a trace model is that if “noise” is assumed to be constant – independent of the memory strength – Weber-law timing requires a logarithmic, hence infinite-range, memory trace. But in fact, the kind of memory trace necessary to explain data on rate-sensitive habituation and Jost’s forgetting law (Chapter 6) is very close to the logarithmic form necessary to account for the Weber’s law and time-discrimination and differentiation data: rapid deceleration at first, followed by slower deceleration. Moreover, as I explain in a moment, the

---

\(^{113}\) Note that some \( \mathcal{V}_M \) values never occur: those corresponding to post-time-marker times longer than the longest interfood interval. Neither reinforcement nor nonreinforcement can be associated with those values.
data suggest that the correct memory-decay function is only approximately logarithmic, although in Weber-law and bisection experiments with moderate stimulus ranges the log and MTS functions are probably not distinguishable. And, finally, the MTS trace is not static, but tuned by the series of training stimuli so that a very large “stimulus” (i.e., interval) range can be captured by a finite process.

The assumption that timing depends on the same mechanism as habituation may also help to solve the credit-assignment problem (i.e., identification of the time marker). The theorist need not assume that the animal “knows” the relevant time marker. Now, every discriminable event will have its own trace, and remembered reinforced values will compete for control of the response. Only good predictors will compete effectively, presumably (see Chapter 10). Moreover, “reset” need not be absolute in this type of model (as it is in cognitive timing theories, absent ad hoc assumptions – see the discussion of reinforcement omission, below). Events preceding the most recent time marker affect $V_M$, hence affect the animal’s “time estimate.” Reinforcement schedules show such effects: for example the postfood response profile shows proactive interference effects on some temporal schedules (e.g., Lejeune, Ferrara, Simons & Wearden, 1997; Staddon, 1974a&b) and the effects of successive reinforcers have been shown to cumulate under some conditions (Horner, Staddon & Lozano, 1997). The fact that events have cumulative effects in the MTS model automatically subsumes timing and counting under the same theory, as I show in a moment.

In the rest of the chapter I apply the MTS model to static phenomena of interval timing: proportional timing, Weber’s law, the reinforcement-magnitude effect, the reinforcement-omission effect, duration-discrimination experiments, and timing and counting experiments. I also describe how some drug effects on timing, usually interpreted in terms of SET, can as easily be fitted into the MTS framework.

**Proportional Timing**

Any variable that changes monotonically with time elapsed since a time marker can serve as an internal clock. Figure 14.2 shows the memory trace after many exposures to a series of stimuli (reinforcers) at 8-time-step intervals. The low points ($R$) indicate the $V_M$ value associated

![Figure 14.3](image-url)
with reinforcement: $V_M^*$. The dashed line shows a possible response threshold. If the threshold is set a fixed distance from the value of $V_M^*$, then the post-time-marker waiting time will be (approximately) a fixed fraction of $t^*$, the to-be-timed interval, so long as $V_M$ is approximately logarithmically related to time. I show in a moment that given the specifics of the MTS model, the assumption that the response threshold is set at $V_M^* + k$, where $k$ is a constant, is sufficient to duplicate both Weber’s law and proportional timing.

Figure 14.3 shows how the MTS memory trace is tuned by the system history. The right-hand panels show the trace right after reinforcement following long exposure to FI schedules with different interstimulus (interreinforcement) intervals, ranging from 5 to 5000 time steps; the left-hand panels show the system initial conditions at the beginning of extinction. There are two points to notice. First, right-hand panels show that the trace following the ISI-5 history declines much more rapidly than following the ISI-5000 history. Second, the left-hand panels show why. These panels show the $V$ values of the ten habituation units at the beginning of the trace. When the ISI is short, fast units early in the cascade retain some strength from stimulus to stimulus, hence block inputs to the later, slower units, which therefore have low or no strength. After the ISI-5 history, for example, only units 1-4 ($V_1 - V_4$) have any strength at the beginning of extinction. Since $V_M$ is just the sum of all the $V$ values, and since each unit decays exponentially at a rate determined by its own time constant, net decay rate is rapid after the ISI-5 history.

Conversely, when the ISI is long, early units discharge completely by the end of the interval, hence fail to block the stimulus signal from reaching later, slower units. At the beginning of extinction, therefore, later units have substantial $V$ values. After the ISI-5000 history, all units are active at the beginning of extinction. The trace consequently decays slowly, since the slow units contribute much of its strength. Conclusion: A ten-unit MTS model can provide a “clock” capable of timing intervals over at least three orders of magnitude.

Figure 14.4 How the Weber fraction changes with interfood interval, according to different trace models. The Weber-fraction estimate (WFE) is $dt/df(t)|$, that is: $1/|slope*IFI|$, where $slope = V_M(t+1) - V_M(t)$. The abscissa shows IFIs from 2-32 time steps. Left panel: Exponential, log and power functions: $f(t) = \exp(-k)$, $f(t) = \ln t$, $f(t) = t^n$. Parameter values are shown in the legend. Right panel: A 3-stage MTS model. Parameter $\lambda$ determines the three rate parameters, $a_i$, according to the relation: $a_i = 1 - \exp(-\lambda i)$ (Staddon & Higa, 1996, Equation 5); parameter $b_i = .1$. (From Staddon & Higa, 1999, Figure 6.)
Weber’s Law (the scalar property)

The Weber fraction, $W$, is defined as the change in physical stimulus value that corresponds to a small fixed change in the value on the psychological scale at that point, divided by the stimulus value. Thus,

$$W(x) = \frac{dx}{x},$$

(14.2)

where $W(x)$ is the Weber fraction at point $x$ on the stimulus scale and $dx$ is the stimulus change necessary to produce a just-noticeable change in the psychological scale (one just-noticeable difference: jnd). Given a particular psychological scale, $z(x)$ — in this case, the trace function — with slope $S(x) = dz/dx$ at point $x$, therefore: $dx = dz/S(x)$. By the jnd assumption, $dz$ is constant. Therefore, we can substitute for $dx$ in Equation 14.2, $dx = k/S(x)$, so that at point $x$, therefore, the Weber fraction is

$$W(x) = \frac{k}{xS(x)},$$

(14.3)

or, in words: the Weber fraction is inversely proportional to slope times stimulus value. I therefore term the quantity $1/|\text{slope} \times \text{value}|$ the Weber fraction estimate (WFE) for a given psychological scale.

In modern terms, Fechner’s insight was that Weber’s law – the constancy of the Weber fraction – follows from two ideas: that sensory dimensions are encoded logarithmically, and that internal “noise” is constant (independent of the encoded value). Logarithmic encoding means that the slope of the internal variable is inversely proportional to its value, so that (the absolute value of) slope multiplied by value is a constant that is inversely proportional to the Weber fraction. Figure 14.4 shows how the Weber-fraction estimate varies as a function of interfood interval for a variety of trace models. The left-hand panel compares the log function (horizontal line) with power and exponential functions. As you can see, the exponential function always violates Weber’s law, because after an initial decrease, the Weber fraction increases with IFI. The Weber-fraction estimate also increases with IFI for power functions, but after an initial rise, further increase is gradual for small exponents ($0 > m > -3$). The right-hand panel shows trace functions from a 3-unit MTS model. For each function, we computed the Weber-fraction estimate at the instant of reinforcement, in the steady state (i.e., after exposure to a number of IFIs). The parameter, $\lambda$, determines how the rate constant, $a_i$, changes from the first unit to the last (see the legend to Figure 14.4). The Weber-fraction estimate increases with IFI for the MTS model, but relatively slowly for $\lambda$ values greater than 1.5, which is the range we have used to model habituation data (Staddon & Higa, 1996). By increasing the number of habituation units (we used three, but physiological data discussed in Chapter 6 suggest that humans, at least, may have ten or more), the constancy of the Weber fraction can be extended to indefinitely large durations. For example, Figure 14.5 shows the 10-unit, ISI-5000 MTS trace from Figure 14.3 replotted with a logarithmic time scale: the line is almost straight, showing that the MTS memory trace over this range is almost perfectly logarithmic.
Figure 14.6 illustrates Weber-law invariance of the MTS model. When steady-state traces at different ISIs (over a 100:1 range) are plotted as a fraction of ISI, and slid down the y-axis so that all pass through the point of reinforcement, \( V_M^* \), the traces superimpose precisely over most of the interval. Consequently, responding guided by a threshold fixed with respect to \( V_M^* \) will show proportional timing. Moreover, because the proportional traces are identical, any threshold variation will necessarily yield response distributions at each ISI value that show Weber’s law and the scalar property.\(^ {114} \) Weber-law invariance is a surprising and powerful property of the MTS timing model.

**Reinforcement-Magnitude (RM) Effect**

A counter-intuitive implication of the MTS theory is that on FI, and similar interval-timing procedures with reinforcement as the time marker, occasional briefer-than-usual reinforcements should reduce postreinforcement pause or break point (“reduce the animal’s subjective estimate of time to reinforcement” in the language of cognitive timing theories). The prediction is counter-intuitive because it implies that an animal will systematically misjudge time-to-food on account of an objectively irrelevant variable: the magnitude of the time marker. This prediction is illustrated in Figure 14.7. The figure shows steady-state memory-strength (\( V_M \)) traces for an eight-time-step schedule. Standard (large) food deliveries are indicated by \( L \), small food deliveries by \( S \). The horizontal line indicates the value of \( V_M^* \), the trace value at the usual time of food delivery. Because the small food deliveries increase \( V_M \) less than the large, \( V_M \) will reach \( V_M^* \) sooner after a small reinforcement than after a large one (arrows). Hence, the animal should begin responding sooner after a small reinforcement than after a large one – and this effect should be immediate.

In experiments that have intermixed reinforcements of different durations, postreinforcement pause is directly related to reinforcement duration, just as this analysis implies (e.g., Hatten & Shull, 1983; Lowe, Davey, & Harzem, 1974; Meltzer & Brahlek, 1970; Staddon, 1970a), and the effect shows up during the first session under the mixed-duration procedure. The data also show, as this model predicts, that the effect is a reduction in pause after the smaller reinforcement durations, rather than an increase after the long.

\(^{114} \) Indeed, Figure 14.6 closely resembles the diagram often used to explain how SET deals with the scalar property: see Figure 13.7 and Church, Meck & Gibbon (1994), Figure 2. The difference is that Figure 14.6 is derived from the theory, not assumed a priori.
This analysis implies certain sequential dependencies. For example, for the sequence in Figure 14.7, if a small reinforcer initiates interval \( N \), and the next interval begins with a large reinforcer, waiting time in interval \( N+1 \) should also be shorter than the average post-\( L \) wait, though not as short as the wait in interval \( N \). Dependencies of this sort have not been studied extensively.

Note, however, that the shorter-reinforcement, shorter-pause effect is not to be expected if all reinforcements are brief. Under those conditions, the horizontal \( V_M^* \) line (i.e., the \( V_M^* \) value usually associated with reinforcement) will simply be lower and postreinforcement pause will be adjusted to that level. In conformity with this prediction, chronic FI experiments with large or small reinforcers show more or less the same postreinforcement pause (Bonem & Crossman, 1988; MacEwen & Killeen, 1991). (The RM effect should, and does, appear to occur during transitions from one reinforcement duration to another, however.)

Reinforcement-Omission Effect (OE)

If a pigeon or a rat is well trained on a fixed-interval schedule, and a reinforcement (R) is occasionally omitted at the end of the interval and replaced by a “neutral” nonreinforcement stimulus (N) such as a timeout or key light change, equal in length to the reinforcement itself, responding in the subsequent interval begins sooner than usual (Figure 14.8). Moreover, this reduction, known as the reinforcement-omission effect (OE), is maintained indefinitely (Staddon & Innis, 1966, 1969). Like the reinforcement-magnitude (RM) effect, it is counter-intuitive, because in all these experiments the time-to-food signaled by reinforcement and by the stimulus presented in its place is always the same. Nevertheless, so long as the omission is not too frequent (no more than 50% of intervals) the omission effect persists indefinitely, in both rats and pigeons (there are some second-order differences between rats and pigeons, however: Staddon & Innis, 1969).

Is reinforcement omission just like reducing reinforcement magnitude? On FI schedules reinforcement suppresses responding (i.e., produces a postreinforcement pause); events similar to reinforcement should suppress it less (stimulus generalization decrement). There is good experimental evidence for the generalization-decrement interpretation. For example, wait time after a “feeder flash” (too-brief operation of the feeder) is less than following reinforcement, but more than after a brief timeout (cf. Kello, 1972; Staddon, 1970b; Staddon, 1972b; Staddon & Innis, 1969). If reinforcement omission is just like a small reinforcer, Figure 14.7 can be used to explain reinforcement omission as well as reinforcement reduction. In both cases, the estimated
time to reinforcement is reduced, the animal begins responding too early in the IFI and response rate across the whole interval is higher than usual (the omission effect).

Figure 14.7 also resembles the “incomplete reset” description for the OE proposed by SET (Gibbon, 1991). The difference between the two accounts is that SET says nothing about why small reinforcers should be less effective time markers than large, whereas time-marker (reinforcer) magnitude is the input variable to the MTS theory. The MTS account is also an advance over the generalization-decrement interpretation, because it suggests why the omission effect is persistent rather than transient.

Nevertheless, there is a problem with the assimilation of reinforcement omission to a reduction in reinforcer magnitude: the reinforcement omission effect depends on the absolute value of the timed interval\(^{115}\). In well-trained pigeons, there is negligible omission effect at short (< 60 s) IFIs, (Starr & Staddon, 1974). Yet the effect shown in Figure 14.7 should occur with any IFI. This dependence on absolute IFI value is one reason that the OE has been interpreted as a sort of proactive memory interference (Staddon, 1974b). The argument goes like this. A short time after the omission stimulus, the not-so-memorable omission stimulus can still exert influence and control waiting time. But at longer intervals, the influence of the prior reinforcement overshadows the influence of the omission stimulus and (because time-since-reinforcement is necessarily longer than the IFI), responding begins too soon, producing the omission effect. (Note the resemblance between this interpretation of the OE and Jost’s memory law: an older, initially weaker memory [for the most recent reinforcer] with time overtakes the newer, initially stronger memory [for the omission stimulus].)

In terms of the MTS model, the proactive interference interpretation requires not one trace but two, one for reinforcements and one for omission stimuli. The two-trace assumption is not unreasonable, since the two events, food and a neutral stimulus, differ in at least two dimensions, not just one as in the reinforcement-magnitude effect. Granted that there are separate traces for N and R, the theoretical questions are: How will the traces differ? And, to which trace should the organism attend; that is, which one will control behavior – or will both exert some influence? There are obviously several ways to answer the second question, but the simplest is just to take the traces at face value: as in the MTS interpretation of Jost’s Law (Staddon, 1997b), let the higher trace be the controlling one\(^{116}\).

The first question, how should the traces differ, is trickier. It will not do to simply make the N-trace weaker, because this will not capture the absolute-time dependence of the omission effect. The assumption of a weaker N-trace also permits the R-trace to be higher than the N-trace even immediately after the N-stimulus under some conditions, which violates our intuition that even a feeble stimulus can be perfectly recalled right after its occurrence. The answer is that

---

\(^{115}\) No experiment seems to have been done to see whether the reinforcement-magnitude effect is in fact independent of interfood interval, however.

\(^{116}\) This assumes that reinforcement and the omission stimulus are equally good predictors. A strong trace that is a poor predictor of reinforcement will presumably be less effective as a controlling stimulus than a weaker trace that is a good predictor.
the N-trace must be not weaker than the R-trace, but *less persistent*. There are several ways to implement this in the model. Perhaps the simplest is to assume that the N- and R-processes have different thresholds late in the integrator chain. If the thresholds for N are higher than for R, N-traces will be less persistent than R traces at long ISIs, because later, slower integrators will contribute less to $V_M$.

The next chapter deals with more complex timing tasks.
Chapter 15
TIME, III

**Bisection and Relative-Duration Discrimination**

Bisection at the geometric mean is consistent with log-like encoding of time — hence with MTS — and inconsistent with scalar expectancy theory (Gibbon, 1981), although a later version of SET, with some additional assumptions, is able to fit these data (Allan & Gibbon, 1991). But SET has not been applied to a very interesting set of data on relative-duration discrimination reported by Stubbs and his associates (e.g., Dreyfus et al., 1988; Stubbs et al., 1994). These data are consistent with a trace model that is not precisely logarithmic. The power function, with small exponent (which is close, but not identical, to a logarithmic function) seems to provide an adequate static account.

A typical procedure in relative-duration-discrimination experiments is for pigeons to peck a center key to produce a red light of one duration that is followed immediately by a green light of another duration. When the green center-key light goes off, two yellow side keys light up. The animals are reinforced with food for pecking the left side-key if the red light was longer, the right side-key if the green light was longer — a relatively simple task that pigeons master easily.

How does discrimination accuracy depend on relative and absolute duration of the stimuli in these experiments? The main effects are that discriminability is directly related to the ratio of the two stimuli, and inversely related to the duration of the second stimulus. For example, average discriminability when the short and long stimuli are in 4:1 ratio is higher than when they are in 2:1 ratio. But if we look at 4:1 pairs when the “4” stimulus is first and the “1” stimulus second, vs. the opposite order, “1” followed by “4”, accuracy is better in the case where the “4” stimulus is first. As Stubbs et al. point out, the second finding is consistent with a memory account, because memory for the first stimulus will be weaker when the second stimulus is the longer of the two. The first finding, of course, follows from log-like temporal encoding. Let’s see how these results and others follow from a trace that is approximately logarithmic.
The trace functions generated by the MTS theory in a given experiment can often be approximated by power function

\[ f(t) = kt^w, \quad w < 0, \]  

(15.1)

where \( k \) is a constant and \( w \) is an exponent that varies from experiment to experiment. In other words, a power function is a static approximation to the dynamic MTS trace. A power function with an exponent on the order of -0.1 is close to logarithmic and was used by Stubbs et al. (1994) and Dreyfus et al. (1988) to analyze their relative-duration discrimination data. I present a slightly expanded version of their account here.

**Response Rule.** The basic idea is that animals use the traces of punctate events — onset and offset of stimuli — as a guide to when to respond in interval-timing experiments. In simple duration-discrimination experiments such as bisection, trace value can be used directly. But in relative-duration-discrimination, no single trace value is a reliable guide; trace values must be combined in some way. The animal must somehow compare (encoded) red and green time intervals and respond differentially depending on the outcome of the comparison. A formal analysis must first identify the type of comparison and then show how it is applied to reduce the comparison problem to a decision on a single decision axis — like \( V_M \) in the simple MTS model — where responding is driven by comparison with a threshold value.

The response rule can be derived by asking, What simple transformation of trace values (of the onset and offset times of the red and green stimuli) at the time of choice best predicts reinforcement on this procedure? The answer, for log-like functions, is that the *difference* between traces provides a good estimate of relative stimulus duration.

Let \( T_R \) and \( T_G \) be the durations of the red and green stimuli on a given trial. The reinforcement contingencies are

\[
\begin{align*}
\text{If } T_R / T_G < 1: & \text{ Response A is reinforced} \\
\text{If } T_R / T_G > 1: & \text{ Response B is reinforced.}
\end{align*}
\]  

(15.2)

To estimate \( T_R \) and \( T_G \), the organism has available traces corresponding to every discriminable change in stimulation. In this case, the relevant events are trial onset, and the offset of the red stimulus (i.e., the time of the red-green transition). The values at time of choice of traces initiated by these events constitute the input to the decision process. Traces for two trial types in which \( T_R + T_G = \text{constant} \) are illustrated in Figure 15.1. Trace A is the trace for trial onset, traces B and B’ are for the time of the red-green transition: B is for the case where the red stimulus is shorter; B’ for the case where it is the longer. C represents the time of choice (green stimulus offset).

At the time of choice (i.e., just after the green stimulus goes off), the value of the trial-onset trace is \( f(T_R + T_G) \) and the value of the red-offset trace is \( f(T_G) \), where \( f \) is the log-like time-encoding function. Perhaps the simplest theoretical assumption we can make is that the organism is able to compute the difference between pairs of traces. This is done for the log and power functions for \( f \) in Figure 15.2, which conveys the same information as the top panel in Figure 6 in Stubbs et al. (1994). Each curve is for a given base duration, from 2 to 24 s. Each base duration is compared with 5 comparison durations in ratios of .25, .5, 1, 2 and 4 – approximately the range studied in the Stubbs et al. experiment. The vertical line at 1 on the abscissa divides pairs in which red is longer than green from those in which green is longer than red. The horizontal dashed line is a possible criterion: above criterion make Response A, below make Response B.

The first thing to notice is that with pure log encoding (right panel), performance is the same at a given red/green ratio, independent of the absolute time values. This is inconsistent with the pattern in the data, which shows differences between long and short absolute durations even if their ratio is the same (Stubbs et al., Figure 6). The pattern in the data is closer to what is predicted by the power form (left panel): at a given duration ratio, performance accuracy – the
slope of the difference function – is always higher for shorter absolute durations. Second, for a reciprocal pair of red:green durations (equidistant from the 1:1 line on the log scale in the figure), such as 1:2 or 2:1, the distance to the criterion (which determines accuracy) is always greater for the 2:1 pair than the 1:2 pair: accuracy is better when the shorter stimulus is second. Because it is also concave upward, like the MTS functions, this prediction is true for the log function also. Third, because the functions with longer base durations are lower, the theory predicts an increasing tendency to make Response B (“red shorter”) at longer absolute durations (when the red-green ratio is held constant), a shift in bias also seen in the data (Dreyfus et al., 1988; Stubbs et al., 1994). Finally, the smaller the red/green ratio (i.e., the longer the second stimulus in relation to the first) the less the effect of base duration; Stubbs et al. report good agreement with data on this point also. The fact that these four very general predictions from a nonlinear time-encoding model agree with these data effectively demolishes the idea that animals encode time linearly.

The Stubbs et al. data do contain one apparent anomaly. In a comparison of broad and restricted ranges of sample times, they report that performance on individual pairs in the restricted-range condition was worse than in the extended-range condition, a result incompatible with any analysis that treats each pair in isolation. They suggest that the longer trial duration under the restricted-range condition may be responsible. I’m not sure what this result means for the MTS model. One possibility is that a dynamic MTS analysis (hard to do in this complex situation) may give predictions that deviate slightly from the static power-function approximation, because any dynamic analysis will carry over some information from one comparison pair to the next. Another possibility is that the differencing rule may be wrong. The obvious alternative is an absolute rule (Dreyfus, Fetterman, Stubbs, & Montello, 1992): to learn to link Responses A and B to specific green durations (trace values). It is likely that learning absolute $V_M$ values is the “default” process, hence easier than learning a differencing rule. But Stubbs et al. found little evidence for control of behavior by absolute stimulus duration. It is nevertheless conceivable that learning the differencing rule is easier given a larger set of exemplars, i.e., in the broad-range

![Figure 15.2](image-url)
condition. Since the differencing rule is better than any absolute rule, this would yield better performance with the broader range. But without some formal model of the process by which the system selects among response rules, this is just conjecture.

The trace-difference analysis explains relative-duration discrimination by means of a single response criterion on a trace-difference decision axis. This process is not sufficient to discriminate among three alternatives, however, which is what is required in the temporal version of the “intermediate-size problem”, i.e., learning to make one of three responses depending on whether an intermediate duration appears first second or last in a sequence of three durations. Given three durations and three traces and a differencing rule, a single criterion is sufficient to identify the longest or the shortest, but a second criterion is required to identify the intermediate duration. Fetterman (1998) has recently reported that pigeons can indeed identify the longest or the shortest in a series of three but fail to perform reliably above chance (or to show transfer to novel durations) in identifying the intermediate duration. Apparently, pigeons are limited to a single criterion.

Overall, the MTS analysis of relative-duration discrimination data emphasizes two important theoretical points: (a) The event trace is log-like, but two kinds of empirical data show that it deviates systematically from the log form: the gradual increase in the Weber fraction at long times; and the effect of absolute time on performance at the same red/green ratio in relative-duration discrimination. Both the power and MTS trace functions satisfy these conditions. (b) The response rule and memory-trace form are to a degree complementary. The trace-difference approach assumes that differencing is a simpler assumption than the forming of ratios and is thus theoretically preferable. Nevertheless, linear time and a ratio response rule gives similar results to a log trace with a differencing response rule. The problem for the linear approach is that deviations from Weber’s law and absolute-time effects in relative-duration-discrimination data require systematic deviations from the log form. Slight deviations from the log form are naturally accommodated within the tuned-trace framework, but not within a linear-time framework.

If pigeons must use a trace-difference response rule, and if the trace is log-like in form, they should be unable to learn to discriminate between two randomly chosen times, \(x\) and \(y\), that differ by a fixed amount \(z\) vs. two randomly chosen times that differ by 0 or 2\(z\). Conversely, animals should have no difficulty discriminating on the basis of duration ratios: e.g., discriminating between randomly chosen times in fixed ratio, \(x/y = k\), vs. ratios of \(.5k\) and \(2k\). The first experiment does not appear to have been done, but Fetterman, Dreyfus and Stubbs (1989) have done a ratio-discrimination experiment. They have shown that pigeons can learn to make one response if two successive durations are in a ratio greater than 2:1 and a different response if the ratio is less than 2:1, for example. Moreover, the pigeons showed much the same accuracy at a range of different criterion ratios, a result perfectly compatible with a single-criterion, log-like-trace model.

---

**Figure 15.3** Mean percentage of correct responses on trials initiated by a short (2 s) or long (10 s) sample at three different choice delays: 0, 10 or 20 s. (From Grant, Spetch & Kelly, 1997, Figure 1.)
Absolute Duration Discrimination and the “Choose-short” Effect

There is an interesting asymmetry in time discrimination that provides very strong support for a decaying-trace “clock”. In the bisection experiment, pigeons and rats can readily learn to make one response following a short-duration (e.g., 2-s) stimulus and another response following a long (e.g., 10-s) stimulus, even if the opportunity to choose is delayed slightly after stimulus offset (e.g., Church, 1980; Spetch, 1987; Spetch & Wilkie, 1982). But if longer choice delays are occasionally introduced, performance worsens dramatically and in a biased way: the frequency of “short” errors increases with delay: the choose-short effect (Grant, Spetch & Kelly, 1997; Spetch & Wilkie, 1982). Figure 15.3 shows typical data from a delayed-matching-to-sample (DMTS) duration-discrimination experiment with 2- and 10-s samples. The upper line shows that as the time-to-choice increases, the percentage of correct responses following the short sample slowly decreases to chance, as expected. But after the long sample, accuracy declines precipitously to below chance at longer delays, because the animal is consistently choosing the “short” response. Why the asymmetry? Why doesn’t choice accuracy following the long sample decline in the same gradual way as it does following the short sample?

The idea that time discrimination is controlled by a trace-difference response rule provides an obvious explanation for the choose-short effect. In DMTS duration-discrimination experiments, the only reliable correlate of stimulus duration is the difference in strength (at the time of the response) between the trace of stimulus offset and the trace of trial onset. Formally, $\delta_S = f(D) - f(T_S + D)$, for the short stimulus and $\delta_L = f(D) - f(T_L + D)$, for the long, where $T_S$ and $T_L$ are the stimulus durations and $D$ is the choice delay. With the power form for $f$, this difference, $\delta$, will obviously always be larger for the long stimulus than the short, $\delta_L > \delta_S$. But if $D$ is suddenly increased, $\delta_L$ will decrease much more than $\delta_S$ and move in the direction of $\delta_S^*$, the reinforced value of $\delta_S$; i.e., the situation will seem more similar to the situation in which a “short” response is appropriate. The argument is illustrated graphically in Figure 15.4.

This analysis predicts the opposite effect if $D$ is decreased, however, because reducing $D$, shifts $\delta_S$ in the direction of $\delta_L^*$, the reinforced value of $\delta_L$. To see this, imagine that $D2$ (rather than $D1$) in Figure 15.4 is the training delay and $D1$ is the testing value. In an experiment with $T_S = 2$ s, $T_L = 8$ s and $D = 10$ s, Spetch (1987) has shown that pigeons perform accurately at the training delay, show a “short” bias when occasional longer delays are introduced, but show...
“choose long” errors when shorter delays are introduced. The symmetry between the “choose short” and “choose long” effects strongly supports the trace analysis.

_The Symmetry Between Timing and Counting_

There is a close resemblance between the discrimination of time and number. For example, Roberts (1997) trained pigeons to make one response following a series of 8 flashes and another after 2 flashes, both series occurring over a 4-s period to eliminate time as a cue. After training, the opportunity to respond was delayed for 2, 5, or 10 s, with dramatic results: responses to the 2-flash stimulus remained mostly correct; but correct responses to the 8-flash stimulus rapidly declined with increasing delay, a “choose small” (or “choose fewer”) effect that parallels the “choose short” effect (see Figure 15.3).

Pacemaker theories explain timing as a sort of counting, which implies obvious similarities between the two and appears to support the pacemaker-accumulator approach. Meck, Church and Gibbon (1985) proposed that the same process might underlie both types of discrimination – an extension of SET that, in its most recent version, allows a single accumulator to count both pacemaker pulses and events. The input to the accumulator is controlled by what Meck terms a “mode switch” which “gates the pulses to the accumulator in one of three different modes depending on the nature of the stimulus, giving this mechanism the ability to act as both a counter and a timer based on the representation of the variable in the accumulator. (Meck, 1997, p. 141-2)” The three modes are “run,” “stop” and “event.” “Run” accumulates pulses from stimulus onset until its end; “stop” accumulates pulses so long as a stimulus is present (i.e., it accumulates across successive stimulus occurrences); and “event” counts events. By suitably combining these modes, the mode-control version of SET can reproduce the similar psychophysical properties of time and number discrimination.

This approach is able to simulate any given experiment, but the theory is vague on the critical matter of how the mode is chosen, on what events and durations are to be counted or timed. The experimenter knows, of course, but how does the model (or the animal)? This is a version of the assignment-of-credit problem, more acute when reinforcement may be contingent on counts or times, or both.

Roberts (1997) explains the choose-fewer effect by adding a “loss” assumption to the mode-control model. He suggests that working memory “loses pulses” at an exponential rate

---

**Figure 15.5** The “choose fewer” effect, simulated by a 3-unit (inset) habituation model. Animals are trained at short response delay _D1_ and set their response criteria accordingly (2- and 8-flash thresholds). When tested at a longer delay, _D2_, trace strength following the 8-flash stimulus shifts in the direction of the criterion for the 2-flash stimulus, producing a “choose fewer” bias.
over time. He goes on to make the argument that as time elapses, the number of pulses in the accumulator after the 8-flash stimulus will approach the criterion for the 2-flash stimulus, hence the increase in “small” responses as the response opportunity is increasingly delayed. He concludes that the data “support the ideas that time and number are represented by the common mechanism of accumulated pulses and that pulses are progressively lost from both time and number accumulators over a retention interval. (pp. 200-201)”

What these data in fact support is simply the idea of memory-trace decay. Nothing in the choose-fewer effect demands pulses or multiple accumulators. Because the MTS model is driven by events and its output accumulates their decaying effects, it automatically explains the resemblances between time and number discrimination. It explains the “choose fewer” effect directly, without any additional assumptions and with none of the elaborate apparatus postulated by the mode-control versions of SET. The argument is trivial. Each flash is an input event and increments the integrator cascade. Two flashes in 4 seconds produces a lower output (trace value) than 8 flashes in 4 seconds. If the criteria for choice are set for a short response delay, then at longer delays, the trace value following 8 flashes will approach the 2-flash criterion, hence responding will increasingly be to the 2-s key: the “choose fewer” effect. The MTS account is illustrated in Figure 15.5, which parallels Figure 15.4.

Because it is real-time, the MTS account makes specific predictions about the effects on choice of the temporal pattern of flashes within the counting period. We don’t know how well these predictions will match the data, since the necessary experiments remain to be done. But obviously many known similarities between timing and counting, such as Weber’s law, proportional counting and the choose-fewer (and, potentially, “choose-larger”) effect, are immediately implied.

**Drug Effects**

Dopaminergic and cholinergic drugs can affect timing behavior. Meck (1983, 1996) has identified two main kinds of effect on a temporal dependent variable (call it $T$) such as peak time in the peak procedure: $T$ can either increase or decrease under the influence of the drug; and these effects can be either both **immediate** ($T$ increases or decreases soon after the drug is administered, with an inverse rebound when drug administration ceases) and **transient** ($T$ returns to the pre-drug level under continued administration, and following the rebound when administration ceases); or the effects can be **gradual** (the change in $T$ occurs over several days), in which case it usually **persists** until the drug is no longer administered, when $T$ slowly returns to the pre-drug level. Interestingly, there seem to be no examples of the other two logical possibilities: immediate, persistent effects or gradual effects that eventually dissipate.

The immediate-transient-type change is usually termed a clock-type effect in SET. It is attributed to a change in “clock speed.” If the clock runs faster under the drug than previously, the animal’s estimate of reinforcer time will be reached too soon: $T$, the temporal dependent variable, will therefore decrease. Dopaminergic agonists such as methamphetamine have this speeding-up effect. Conversely, dopaminergic antagonists such as haloperidol have the opposite effect: clock speed is decreased so that $T$ increases.

The gradual-persistent type of effect is usually termed a (reference) memory-type change in SET. The interpretation here is that “memory-storage speed” is altered by the drug. There is no immediate change in performance “because the vast majority of time values stored in reference memory were accumulated during baseline training under conditions of normal memory-storage speed, not one distorted by drug administration. (Meck, 1996, p. 237)” As the distorted memories accumulate during training under the drug, $T$ shifts appropriately. The cholinergic antagonist physostigmine produces a slow decrease in $T$, whereas the agonist atropine produces a slow increase in $T$. 

201
MTS theory provides alternative interpretations for these effects. The immediate-transient change resembles the reinforcement-magnitude effects discussed earlier. Reducing reinforcement magnitude, for example, produces a transient decrease in $T$; increasing reinforcement magnitude produces a transient increase. Thus, one alternative interpretation of so-called clock-speed effects is that they cause an immediate change in the memorability of the time marker. Tests of drugs that produce these effects with a variety of different time markers may therefore produce quantitatively different effects on $T$: a given drug and dose may produce a large effect with a not-very-salient time-marker like a neutral timeout, and much smaller effect with a highly salient time-marker like a large reinforcer. It is also possible that immediate-transient changes reflect changes in the time constants of the integrator chain. This interpretation is close to the SET clock-speed-change interpretation of these effects. In this case, differences between different time markers should be minimal. Quantitative experimental data, showing the effects of different dopaminergic drugs at different $T$ values could help decide this issue.

Another prediction is that other drugs known to affect short-term memory (e.g., in delayed matching-to-sample experiments) should also produce immediate-transient effects on time discrimination.

The gradual-persistent effect is much the same as if the to-be-timed interval were to be changed. The change in $T$ here is also usually (but see Higa & Staddon, 1997) gradual, but persists so long as the new interval is maintained. In the MTS model a gradual-persistent effect would be interpreted as a change in the remembered value of $V_M^*$, $V_M^*$, the trace value associated with reinforcement. Under normal conditions, $V_M^*$ is approximately equal to $V_M^* - $ which is why the relation between $T$ and the actual to-be-timed interval is usually a power function with unit slope, i.e., linear (see the discussion of power functions and log time in Chapter 13). If (long-term) memory is imperfect, however, the two values may drift apart, so that at the time of testing, the animal may show a too-long or too-short value of $T$, depending on the direction of drift. This interpretation suggests that the amount of “memory shift” under a given drug regimen may depend on the time between experimental sessions, i.e., the time allowed for drift: if sessions are closely spaced there may be less drift than if they are 24 or 48 hours apart. It will also be interesting to look at the popa clock-wer-function exponent when steady-state $T$-values under drug and no-drug conditions are compared.

Figure 15.6 Temporal tracking on a response-initiated-delay schedule (see Figure 3.11) in which successive intervals follow a sinusoidal pattern (smooth light line). Light lines with markers are data for individual pigeons; heavy line is the average. Note that waiting time tracks the input interval series with a lag of about one interval. (From Staddon & Higa, 1997, Figure 2.)

$N$ depends on the duration of preceding interfood interval $N-1$ (Higa & Staddon, 1997; Innis & Staddon, 1971; Lejeune et al., 1998), might also shed light on these two hypotheses. On periodic schedules, according to one version of SET, reference memory contains many values of the time-of-reinforcement, $t^*$, which is why cholinergic drugs, which distort the translation from working
to reference memory, take some time to have an effect on temporal performance (gradual-persistent effects) – “because the vast majority of time values stored in reference memory were accumulated during baseline training under conditions of normal memory-storage speed, not one distorted by drug administration.” On cyclic schedules, however, performance is strongly determined by the preceding interfood interval (see Figure 15.6 for an example), which implies that reference memory under these conditions contains few $i^*$ values. Consequently, on cyclic schedules cholinergic drugs should have their effect within a few intervals, rather than taking several sessions. MTS theory seems to make a similar prediction, albeit from different premises. If we interpret gradual-persistent effects as drift in $V^*$ then the time for drift on cyclic schedules is much less than on regular FI schedules, hence the effect of the drug should be less.

TIMING: CONCLUSION

The ability to discriminate between stimuli of different durations and to respond differentially at times associated with the presence or absence of reinforcement is widespread in the animal kingdom. The clock-like properties of this behavior, which can be much improved with suitable training, has made the inference of a real “internal clock” almost irresistible. There are many kinds of clock, but the best ones operate by counting pulses. Unfortunately, the very power of pacemaker-accumulator (PA) clocks disqualifies them as models for interval timing in animals. The longer the interval (and the larger the number of accumulated counts), the more accurate PA clocks become. Animal timing is quite different: relative accuracy is essentially constant over a limited range of times, and actually declines over longer ranges. The Weber-law property can be reconciled with “Poisson variability” in the pacemaker only by additional assumptions: that rate variation occurs only between, and not within trials; or by assuming such a fast pacemaker that other, “scalar” sources of variance make the major contribution and the PA feature becomes irrelevant. In either case, the real theoretical work is being done by something other than the PA mechanism, because of its fundamental incompatibility with Weber’s law.

What is left in scalar expectancy theory, once the PA property has been eliminated, is the comparison process, and the assumption that time is encoded linearly, with scalar variance. I have accepted the comparison idea as a working hypothesis in this chapter and the preceding one. But I believe that the “time-left” experiments, which seem to support linear (rather than log-like) encoding of time, are open to alternative interpretations. Moreover, there are several lines of argument that favor log-like encoding: bisection data, a theoretical argument that derives power-function empirical relations from log-like encoding of working and reference memory, and data on absolute and relative temporal discrimination. The choose-short effect provides very strong evidence for a trace interpretation of time discrimination and is incompatible with PA-type models, absent auxiliary assumptions. Other data that appear to support SET – start and stop distributions, start-stop and start-spread correlations, in the peak procedure, for example – only partially confirm the simple version of SET, can only be partially accommodated via extra assumptions (thresholds, variation in reference memory), and can be explained in other ways. Drug effects compatible with SET are also compatible with the MTS approach. Both approaches seem to make some sense physiologically, so that physiological data do not discriminate between them.

What remains by way of direct support for the PA account are quantitative matches between empirical distributions of peak times and related measures and predictions from versions of SET (e.g., Church & Gibbon, 1982). The most striking fact about these distributions (in peak-procedure experiments, for example) is that they are more or less symmetrical, rather than skewed, as would be expected if the response distribution were directly derived from normal variation on a log-like internal code. I have not discussed these distributions in detail for two reasons, one general, the other specific. The specific reason is suggested by Platt’s (1979) comment “how can the relationship between [a] stimulus and its internal representation be inferred from
responding which may be the result of additional transformations required by the reinforcement contingencies?” That is, granted we know that the distribution of responses in time in procedures like the peak procedure can be molded by varying the probability of reinforcement and nonreinforcement at different times, how can we — without knowing the laws that govern the effects of reinforcement — settle on a particular distribution as “the” correct one? (See the discussion of Figures 13.1 and 13.2 earlier.)

The general reason for deferring discussion of distribution form is that Gaussian-type distributions commonly appear whenever repetitive observations are averaged or when dealing with a multistage process. Both these conditions are fulfilled by temporal discrimination experiments. The usual data in peak-procedure experiments are averages across hundreds of trials and often ten or more animals. The preferred measure for demonstrating timing functions is peak time (i.e. the peak of the response-rate vs. time distribution – the average of the wait-run-stop pattern generated on each empty trial), which is statistically guaranteed to be more symmetrical than the distributions of start and stop times that underlie it. As Figure 14.6 shows, the MTS approach readily predicts both Weber’s law and right-skewed start-time distributions. It can also generate symmetrical peak-time distributions (Staddon, Higa, & Chelaru, 1999).

There are better ways to test behavioral theories than through precise correspondence between Gaussian predictions and highly averaged data. It is better to look first at individual animals and data from a wide range of timing situations, and begin with qualitative (rather than quantitative) predictions from a theory that uses as few a priori theoretical concepts as possible. Detailed quantitative predictions may be appropriate once a qualitative match between theory and a broad range of data gives us confidence that the major assumptions are correct. The field is not yet at that stage, particularly as far as assumptions about learning (associative) mechanisms are concerned. Consequently, even though the ultimate objective is to come up with a fully dynamic, real-time timing theory, my focus has been on static timing phenomena and effects that can be explained without assuming any but the most general learning principles.

I have argued in these two chapters that a multiple-time-scale (MTS) memory model, derived from ubiquitous data on rate-sensitive habituation – and consistent with well-established quantitative memory principles – can account for the major features of data from a wide variety of time-related experiments: proportional and Weber-law temporal discrimination, transient as well as persistent effects of reinforcement omission and reinforcement magnitude, bisection, the discrimination of relative as well as absolute duration, the choose-short effect and its analogue in counting experiments. The many resemblances between timing and counting are an automatic consequence of the MTS model. Absolute-duration effects in experiments on the discrimination of relative duration show that the static form of trace left by a time marker is only approximately logarithmic, hence better fit by the MTS model than by a logarithmic model.

The major uncertainties in the MTS approach are to do with learning and long-term (associative) memory. How are trace values “learned”? When, and how, are trace values “stored” and “retrieved”? – if indeed these metaphors are even appropriate. Only when the learning issue is settled can we begin to apply the theory to experiments that deal with expectations – choice between delayed reinforcers, the time-left procedure, etc.; and with true dynamic effects, such as the rapid timing effects now widely confirmed (e.g., Wynne & Staddon, 1988; Lejeune et al., 1997) as well as older data on cyclic temporal schedules (e.g., Staddon & Higa, 1991; Staddon, 1969). These rich and orderly data, demonstrable in individual animals, are still not understood. They pose a real challenge for future developments in the theory of interval timing.
Chapter 16

AFTERWORD

Our discussion will be adequate if it has as much clearness as the subject-matter admits of, for precision is not to be looked for alike in all discussion...it is the mark of a wise man to look for precision in each class of things just so far as the nature of the subject admits; it is equally foolish to accept probable reasoning from a mathematician as to demand from a rhetorician scientific proofs.117

This book is an argument for a way of doing behavioral science that is different from the current mode. Rather than emphasizing the testing of hypotheses or the collection of data for its own sake, I have focused on induction. Simple real-time, albeit often qualitative, theoretical accounts that describe a wide range of phenomena, from simple orientation behavior to habituation to feeding and time discrimination, may prove more useful than precise experimental tests of models that lack any inductive base.

Psychology is part of biology. Darwinian evolution is the greatest biological discovery of the past two centuries. Yet this discovery depended on experimentation almost not at all. Darwin was mainly concerned to explain a vast array of facts about the natural world. He did some simple experiments, on the survival of seeds in sea water, on the physical properties of orchid pollination, on bee behavior and several others. None of them depended on statistics or group averaging, and none played a major role in either the discovery or the acceptance of his theory. The real basis for the theory was its ability to explain and relate.

Even modern molecular biology depended in its origins much more on speculative intuition than formal hypothesis testing. Mendel did indeed do experiments; but his experiments were designed to generate hypotheses, not to confirm them: “The object of the experiment was to observe these variations in the case of each pair of differentiating characters, and to deduce the law according to which they appear in successive generations.” (Mendel, 1865/2000) The experiments were to generate data that might allow him to infer a law, not to test a law already known: “experiments to find,” not “experiments to prove.” And Watson and Crick notoriously operated by guess and by golly – arriving at their model of DNA through appropriating (sometimes by dubious means) facts established by others combined with free speculation (Watson, 1970). The testing of explicit hypotheses has played a minor role in the development of biological understanding.

Unfortunately, much behavioral research has degenerated into what Richard Feynman once called “cargo-cult” science. It follows the procedures of science – careful experimentation, quantitative data, the latest statistical tests. But it lacks the soul of science, which is simply curiosity, the desire not just to “get results” but to understand how the world works.

Too much is made in psychology of hypothesis testing, especially when the subject matter is groups rather than individuals (Ioannidis, 2005; Staddon, 2019). It is no accident that the old name for the “natural science” (as opposed to “social science”) approach in psychology is “experimental psychology.” Why argue against experimental test, which has become so basic to behavioral science? Not because testing is unimportant to science in general. But because it is much less important to psychological science than people think; and because as it is actually done, it rarely advances our understanding.

Psychological experiments are often less conclusive than they appear. First, they too often involve group averages, which can give no information about individual dynamic processes. Second, because of the almost universal use of inferential statistics, most experiments provide little more than one bit of information: “X is greater (or less) than Y, hence the hypothesis is supported (or not).”

More quantitative studies are of two kinds. The informative ones tend to be curiosity driven: what is the function that relates variable X to variable Y? Or, does X have an effect on Y? These are experiments-to-find. The less informative tend to restrict testing to a narrow range of conditions, the parameters of which rarely form part of the relevant theory. The famous “matching law” of choice, for example, is found only under strictly defined conditions — a changeover delay of a certain magnitude, for example, which is not part of the theory. Cognitive timing theories are always tested in conditions with strong time markers, allowing the theories to ignore the role of memory for the time marker. Experimental results often cannot be generalized beyond the particular experimental arrangement that produced them.

Of course, any theory must eventually be tested. The questions are when, and how. As for “when,” I contend that a hypothesis is not worth testing until it has some inductive base. That is, until it has been shown to explain a range of existing data: “The line of argument often pursued throughout my theory is to establish a point as a probability by induction and to apply it as hypotheses to other points and see whether it will solve them.” (Darwin)

As to “how,” I describe several experimental tests of the timing model discussed in Chapters 14 and 15. The general issue of how to test a partial model is taken up in Chapter 9 in connection with the complex Morato et al. (1995) feeding experiment.

How does one test a regulatory model that we know describes only part of what the animal is doing? The usual answer is “Change the preparation so as to exclude extraneous variables.” This is the “spinal preparation” solution, but unfortunately it is usually impossible. We cannot easily interfere with an organism in such a way as to disable associative functions leaving only regulatory ones, for example. What we can do is subject the intact animal to a variety of treatments and see how much of the resulting behavior falls within the patterns prescribed by theory, recognizing that a perfect, quantitative match is unlikely. There are no statistical techniques for this approach and no guarantee that we will arrive at the correct conclusion. But that isn’t necessarily a problem with the method. It is just reality. (Indeed, any method that guarantees a correct answer is self-condemned as a fraud, because that is not how science works.)

So, how to proceed? I suggest that the place to start is with the dynamics of behavior, the pattern of recurrent physical acts in real time and the way that this pattern is affected by the temporal and spatial pattern of physical stimuli. My emphasis has been behavior that can be studied in animals as well as humans; and on behavior that is learned — in the sense that it is modifiable by experience — but not associative. Why? Because non-associative behavior is simpler to understand, and because it is intimately tied up with associative learning, which probably cannot be understood until we first partition out the non-associative bits. Why animals? Because they are simpler (if only because they do not speak115); because ethics permit us to do better and longer-term experiments with animals than with people; and because evolution tells us that human and animal psychology are likely to involve common principles. Indeed, one of my themes has been that habituation and interval timing in animals involve processes known to be involved in human memory since the time of Ebbinghaus.

Another theme in this book is parsimony: It is indeed vain to do with more what can be done with less. Parsimony is not just a desirable feature, like a sunroof or heated seats. It is what gives scientific explanation its power. Who could object? Well, some cognitive scientists seem to: Johnson-Laird, for example, has written: “[T]he sole purpose of science is not to frame parsimonious laws…science aims to explain phenomena, not merely to describe them in laws. And explanations, of course, take the form of theories.” (1988, pp. 17-18). I’m not sure exactly how Johnson-Laird means to distinguish between laws and theories, but he seems to imply that

---

118 Students of comparative cognition will object that recent experiments show some (feeble) linguistic ability in apes and parrots. Yes, but the gulf between human and animal language remains vast.
theories should be judged by criteria other than parsimony, and here I must disagree. I defy anyone to point to an accepted theory in science that is less parsimonious than the theory it replaced.

The response to this challenge might well be to raise another objection: how should parsimony be defined? This is a problem that philosophers of science have wrangled over for decades. Part of the difficulty comes from the fact that parsimony is a ratio, of facts explained divided by assumptions made. And neither assumptions nor facts can be easily quantified. Nevertheless, just because a property cannot be defined easily does not mean that it doesn’t exist. Health, beauty and intelligence are also hard to define, but none doubts that some have more of them than others.

Theoretical behaviorism is, well, theoretical. So, (pace Johnson-Laird) is cognitive psychology. So, what’s the difference? The difference is parsimony. Look, for example, at Figure 16.1, which shows block diagrams of two similar cognitive theories of interval timing: Gibbon’s SET (top) and Gallistel’s version of it. Gallistel asserts that “there is no way of distinguishing between [the two models] by purely behavioral experiments.” (1990 p. 309) This already raises questions, but more critical from my point of view is the large number of essentially ad hoc components in each model. How necessary are all these elements? How well-specified are they? And how can they be independently established? Take “Memory for individual intervals” (top), for example. How big is this memory? Are all intervals represented equally, so that an interval experienced 50 minutes ago is as influential as one experienced 2 minutes ago? (I hope not, because data show that recency is important: Wynne & Staddon, 1992.) What about the “time of occurrence file” (bottom): How big is it? What are its dynamics? That is, when is it updated and how? The upper model has a pacemaker, whose properties are in fact irrelevant to the model. How many other aspects are redundant? The lower has a “clock”: what kind of a clock is it? How big a data set is to be explained by these models? It seems pretty clear that they cannot deal with all the data on cyclic, progressive and “impulse” schedules (indeed, these pose a problem for all timing theories). How many parameters are involved in these models (at least five, plus whatever is needed for the interval memory, presumably). How are these chosen?
I don’t want to be unfair — these sorts of questions can be raised about any model. But the complexity of these cognitive timing models, and the lack of independent ways to verify their elements, makes them especially vulnerable. On the other hand, as I tried to explain in Chapters 14 and 15, the elements of the (admittedly incomplete) timing model I described can be independently verified and I have tried to eliminate anything that does not contribute to the explanatory power of the model.

The three obligations of theory are to eliminate unnecessary assumptions; if possible, to verify independently the elements of the model; and to extend the model to as wide a range of data as possible. Cognitive theories emphasize none of these aspects. Instead, their focus is on plausibility — mentalistic plausibility or a priori plausible notions like a ticking clock — and on precise fit of quantitative experimental predictions. But plausibility has a bad track record as a guide to truth — think of the flat earth, the geocentric universe and fixity of species, all highly plausible notions in their day. I also contend that precision is not to be expected from models that can never capture all the processes involved in (for example) interval timing. If precision is nevertheless achieved, it is often done at the cost of restricting the experimental situation in such a way that the fit of a complex model to relatively constrained data is almost guaranteed.

What is the alternative? It is notoriously easier to find fault than offer something better. The problem is that theoretical science is not to be done by formula. There is no algorithm for discovery and precision per se is no guarantee of truth. There are no hard and fast rules. Plato pointed out that difficult ideas can only be conveyed through examples. This is what I have tried to do: to show by example that a number of phenomena hitherto explained in cognitive ways — trial-and-error learning, ecological adaptations of feeding behavior, short-term memory, spatial navigation and interval timing — can be explained by simple mechanisms that have nothing “cognitive” about them.

Consider again the problem of feeding regulation. Clearly, feeding is complex process that involves both regulatory factors and associative learning. Both are involved in every experiment. How then are we to understand either process, absent some physiological means of blocking one or the other? In Chapters 7-9, I tackled this problem by showing that a simple regulatory process behaves in many ways like a real rat when subjected to a variety of regulatory challenges: interruption of feeding, increases or decreases in diet richness, work requirements, etc. The point is to show a pattern of qualitative similarities between the behavior of the animal and of the model. If no other model does better, this one can be accepted as an accurate picture of one aspect of the feeding system. What I offer is a precise (hypothetical) mechanism and qualitative test. What psychology has frequently provided is vague assumptions and (spuriously) exact test.

My contention is that our understanding of learning mechanisms advances through the construction and qualitative comparison of parsimonious models that capture more and more of the animal’s behavior. My method of argument is just to show how this might be done in a number of specific instances, a few of which are the topic of this book.


Spetch, M. L., & Wilkie, D. M. (1981) Duration discrimination is better with food access as the signal than with light as the signal. *Learning and Motivation, 12*, 40-64.


https://dukespace.lib.duke.edu/dspace/handle/10161/6003


Staddon, J. (2017) *Scientific Method: How science works, fails to work or pretends to work*. Taylor and Francis.


Adaptive Dynamics

Staddon


Watson, J. B. (1913) Psychology as the behaviorist views it. Psychological Review, 20, 158-177.


FIN