

A Dynamic Route Finder for the Cognitive Map

Alliston K. Reid
Wofford College

J. E. R. Staddon
Duke University and University of Western Australia

Cognitive behaviorist E. C. Tolman (1932) proposed many years ago that rats and men navigate with the aid of cognitive maps, but his theory was incomplete. Critic E. R. Guthrie (1935) pointed out that Tolman's maps lack a rule for action, a route finder. We show that a dynamic model for stimulus generalization based on an elementary diffusion process can reproduce the qualitative properties of spatial orientation in animals: area-restricted search in the open field, finding shortcuts, barrier learning (the Umweg problem), spatial "insight" in mazes, and radial maze behavior. The model provides a behavioristic reader for Tolman's cognitive map.

The cognitive behaviorist Edward Tolman spent much of his career devising clever experiments to show that stimulus–response accounts of rat behavior cannot be correct. Some of his most striking demonstrations involve spatial learning. One such example is shown in Figure 1, which depicts 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 box. 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 that they do not know Paths 1 and 2 share a common, blocked, segment? Tolman and Honzik's rats behaved intelligently and usually 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 (1935) famous objection that Tolman's theory left the rat "buried in thought" (p. 143) at each choice point (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 does not discuss barriers, open-field navigation, or cognitive maps. More recently, Hampson (1990) has set out an ambitious connectionist program intended to model Tolman's maze results and many others in the animal learning literature. His model, "which is actually more a collection of components than a single monolithic structure, traces a path from relatively low-level neural/connectionist structures and processes to high-level animal/artificial intelligence behaviors" (p. 1). We mention points of contact between Hampson's approach and ours as we develop our much simpler route-finding model, which is based on the familiar behavioristic idea of stimulus generalization. We identify a class of spatial problems, including the Tolman and Honzik problem, that can be solved by a very simple local process, operating on 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 noncognitive route finder. Our intention is not so much to provide a comprehensive theory of spatial orientation as to show that the processing requirements for behavior that shows spatial "insight" are simpler than previously thought. The model we describe is not complete, but may be an essential ingredient in any comprehensive model for spatial orientation in animals.

We first describe the model and illustrate how it simulates the main features of stimulus generalization. We go on to show how the same process can simulate the behavior of animals in

Alliston K. Reid, Department of Psychology, Wofford College; J. E. R. Staddon, Department of Psychology: Experimental, Duke University, and Department of Psychology, University of Western Australia, Nedlands, Australia.

Research support was provided by grants to Duke University from the National Science Foundation and the National Institute of Mental Health and grants from Eastern Oregon University and Wofford College. We thank Nestor Schmajuk for insightful discussions of many of the topics addressed in this article.

Correspondence concerning this article should be addressed to Alliston K. Reid, Department of Psychology, Wofford College, Spartanburg, South Carolina 29303 or to J. E. R. Staddon, Department of Psychology: Experimental, Duke University, Durham, North Carolina 27708. Electronic mail may be sent to reidak@wofford.edu or staddon@psych.duke.edu.

a variety of spatial situations. We conclude by pointing to some limitations on the simple model.

A Local-Process Model

Our route finder is based on a diffusion model for the dynamics of stimulus generalization (Shepard, 1958; Staddon & Higa, 1991; Staddon & Reid, 1990). Some kind of diffusion or spreading activation is a common element in theories of adaptive behavior, beginning with Pavlov (1927); although our model, unlike Pavlov's, is intended to explain behavior rather than brain physiology. We show in the next section that a minimal dynamic generalization model is able to account for a number of properties of spatial orientation, including properties often thought to demand "insight." But first we describe one- and two-dimensional versions of the model.

One-Dimensional Model

The one-dimensional diffusion model assumes that a single stimulus dimension can be represented as a line of units (cf. Figure 2, inset). Each unit has connections to its immediate neighbors (this assumption is relaxed in a moment). Associated with each unit is an activation strength, V_i , which we term reward *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 process analogous to heat diffusion; 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 1, below).

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*

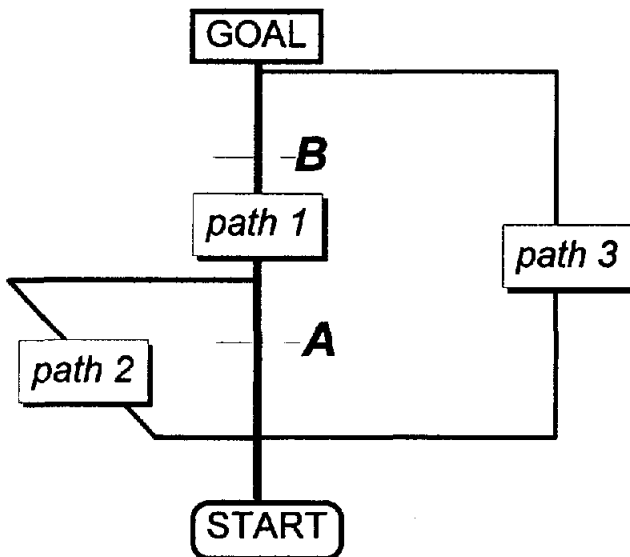


Figure 1. The maze apparatus used by Tolman and Honzik (1930). Path 1 is the shortest path from start box to food box, Path 2 the next shortest, and Path 3 the longest. Blocks were placed at Points A and B during different phases of the experiment.

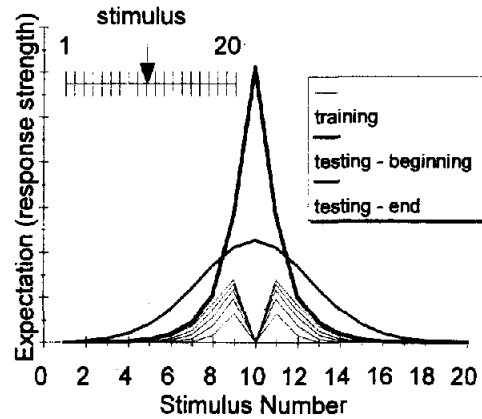


Figure 2. Stimulus generalization according to the diffusion model. The inset represents 20 stimulus-representation units, each with bidirectional connections to nearest neighbors. Each unit has an activation value that changes according to Equation 1 and the other assumptions in the text. The graph shows effect on the activation profile of training with reinforcement delivered in the presence of Stimulus 10 during Time Steps 1, 5, . . . , 25. Fine lines = activation profile during (nonreinforced) Time Steps 4, 8, . . . , 24; heavy line = activation profile at Time Step 30; intermediate line = activation profile at Time Step 50.

units. In the case of spatial generalization, one and only one unit will be active at any instant ("stimulus" in the figure) — 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, the value is set to zero.

Formally, in discrete-time notation,¹ $V_i(t)$ is the expectation associated with unit i at time step t . At each time step the expectation of each unit moves toward 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

$$\Delta V_i = \alpha[(V_{i+1} - V_i) + (V_{i-1} - V_i)], \quad 0 < \alpha < 0.5, \quad (1)$$

where ΔV_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. The α is a diffusion-rate parameter: Larger values of α produce faster diffusion. Equation 1 is a discrete-time version of the standard diffusion equation (e.g., Banks, 1994; Berg, 1983).

The effects of stimuli and of reward and nonreward are incorporated via two assumptions:

Activity. A unit is either *active* or *inactive*, depending on whether or not its stimulus is present (and attended to). If its

¹ Because of the nonlinear reinforcement/nonreinforcement rule, and the nonlinear environment-interaction rule, predictions from the model can only be obtained by simulation. We therefore use discrete-time (rather than continuous-time) formalism, which translates directly into computer implementation.

stimulus is absent, the unit is inactive and the expectation variable changes according to diffusion (Equation 1). If its stimulus is present, the unit is active and the expectation variable is set to a fixed value, which is determined by whether reward or nonreward occurs during that time step.

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 operant response is made to it. If the appropriate stimulus is absent, or if it is present but not responded to (not attended to, in cognitive parlance), then the unit is treated as inactive.

Reward. Reward events set $V_i(t)$ for the active unit to a value directly related to the reward magnitude, $S(t)$. ($S[i] = 1$ in all our simulations). If reward is present, reward expectation equals reward value; if no reward is present, reward expectation equals zero.

These two assumptions allow the expectation surface to change in real time as a function of the temporal patterns of reward and nonreward in the presence of different stimuli. When we generalize the scheme to two dimensions, we will 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 2, 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 reward is presented at $t = 1, 5, 9$, etc. (i.e., every fourth 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 nonrewarded time steps; on rewarded time steps, $V_{10} = 1$). When reward 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 reward 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 perceived similarity of one physical stimulus to another; and reinforcement-expectancy processes that determine how the animal should respond to stimuli more or less similar to the rewarded stimulus. We represent the perceptual processes by a single dimension (line of units) along which stimuli are varied. The diffusion process models the expectation process 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 2. In the first phase, an animal such as a pigeon is trained with a single stimulus (usually termed 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 experimenter records the frequency of responding in the presence of each of these stimuli. 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 versus wavelength (say), a generalization gradient, resembling one or other of the “testing” curves in Figure 2.

Peak shift. One of the more reliable phenomena of generalization, as it is measured in these experiments, is peak shift (Hanson, 1959). To demonstrate peak shift, the first phase of the generalization experiment 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−. If S+ is 550 nm and S− is 540 nm, the new peak might be at 560 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 (cf. Spence, 1937; Staddon, 1977).

Figure 3 shows a simple simulation of peak shift with the diffusion model. S+ (Stimulus 10) and S− (Stimulus 9) are alternated at intervals of four time steps 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 dem-

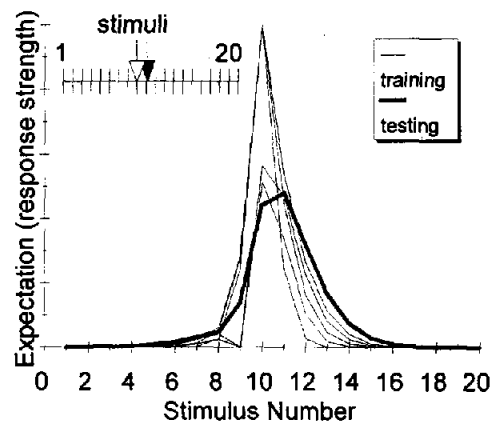


Figure 3. Generalization peak shift simulated by the diffusion model. Stimuli 9 and 10 alternated every 4 time steps for a total of 3 alternations (24 time steps). Reinforcement was always present during Stimulus 10 (S+), never during Stimulus 9 (S−). Fine lines = activation profile during training (see Figure 2); heavy line = activation profile 2 time steps after training. Note that the peak is shifted from S+, away from S−.

onstration shows that the inhibitory assumption is not necessary for peak shift).

The diffusion model is compatible with Terrace's (1964) observation that "errorless" discrimination training does not produce peak shift. In the model, peak shift occurs because reward 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 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), although we do not show these simulations here.

In our simulations, we have assumed 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, but the steepness of the gradient does not seem to increase over extinction trials. If anything, the opposite trend is true: The gradients appear to flatten with the passage of trials. Also in this study, the shape of the gradient is difficult to specify. On the right-hand side of the second test series, for example, the form appears to be ogival, whereas by the fourth series both sides of the curve appear to be somewhat convex downward. (pp. 21-22)

Nevertheless, the flattening effect has not always been observed (e.g., Hoffman & Flesher, 1961), 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 (Rubin & Wenzel, 1996; Staddon, 1998). Hence, a more realistic (but also more complex) version of our 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 multistimulus effects, as the simple model. It is important to emphasize that none of the qualitative predictions we discuss depend on this aspect of the model, so we feel reasonably comfortable about leaving it for future development.

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 feeding at the presumptive new location [von Frisch, 1967]. (Gould & Gould, 1982, p. 281)

Gould and Gould termed this extrapolation "slightly eerie." We have not been able to find a detailed account of this phenomenon, but Figure 4 shows how something of this sort arises naturally from the diffusion model. The inset shows the spatiotemporal pattern of rewarded and unrewarded locations experienced by bees in a study such as von Frisch's. On Day 1 they are fed at close-by source F1 (first filled circle in the inset). On the second day, no food is available at F1 (gray circle below F1 in the inset), but after a delay shorter than a day they find food at the more distant source, F2 (second filled circle in the inset). This pattern is repeated on a 3rd day, with nonreward at F1 and F2 and reward at F3. On the 4th test day, the bees find no food at any of the three locations. The graph shows the expectation profile as a function of location (distance from the hive) after the model has spent some time searching (i.e., with no active units) after failing to find reward at any of the previous three locations. The peak activation is displaced away from hive, at the fourth location, which has never been rewarded. Obviously, the exact form of expectation profile at different times depends on the details of the bees' search pattern, the time and distance between food sites, and so on and these details are not available. Nevertheless, Figure 4 illustrates the fact that a simple form of spatial extrapolation is a robust property of this kind of model.

Two-Dimensional Model: Spatial Orientation

Spatial 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

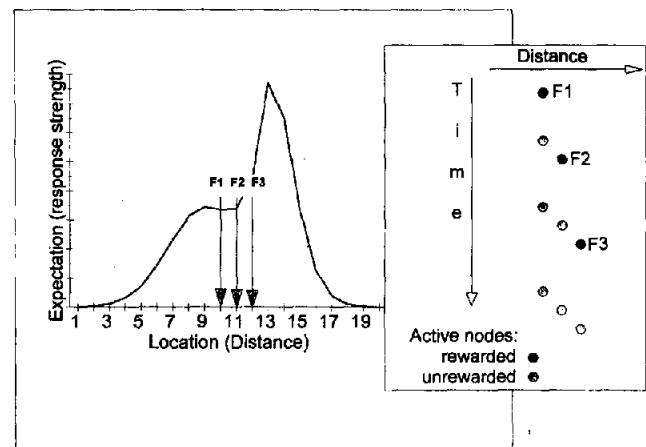


Figure 4. Spatial extrapolation. The inset represents spatiotemporal pattern of reward and nonreward in an experiment by von Frisch in which honeybees were fed each day at a point farther and farther from the hive. F1 = food site on Day 1; F2 = food site on Day 2, and so forth. Filled circles are rewarded sites; gray circles are visited but unrewarded sites. The graph shows the expectation profile versus distance after unrewarded visits to all three sites on Day 4. See text for details.

(within some spatial resolution), that is, a many-one mapping of spatial location on to map node. 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 an orientation 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 later discussion), we are not concerned with *how* the animal knows where it is: Any process—landmark learning, dead reckoning (path integration), viewpoint memory, GPS,² etc.—that can provide the required mapping of position on to state will do (see Gallistel, 1990; Schmajuk & Blair, 1993, for discussions of these issues).

The simplest map is shown on the left of Figure 5. 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 later examples) because it contains no information about adjacency and because no information can pass from one state to another. The other three examples in the figure show different types of map connectivity. In our simulations, we used eight-neighbor connectivity, except for restrictions due to barriers or the edges of the map.

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

$$\Delta V_i = (\alpha/N_i) \sum_{j=1}^N \kappa_{ij}(V_j - V_i), \quad j \neq i, \quad 0 < \alpha < 1, \quad (2)$$

where α is the diffusion-rate parameter, N_i is the number of units connected to unit i , and κ_{ij} is a parameter that is equal to 1 when unit j is connected to unit i and 0 otherwise.³

The only difference between Equations 1 and 2 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. If the underlying map is a rectangular lattice (cf. Figure 5), each element has four neighbors ($N = 4$). In the simulations that follow, we use an eight-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 reward assumptions are the same as in the one-dimensional case. For the two-dimensional case, we need to add an assumption about barriers, as well as an action rule.

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 κ_{ij} is set to zero and N_i is reduced by 1. 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 our 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 our reward rule and the dynamic nature of the expectation surface. If no reward is available at a given point, V_i is set to zero, which ensures that adjacent points will generally have higher V values. If reward 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 condi-

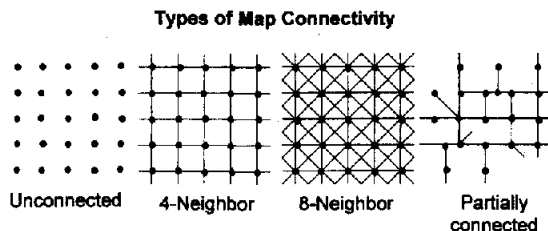


Figure 5. Two-dimensional map structures. In the *unconnected* map, there is a many-one mapping of environment onto state (states are indicated by the filled circles), but no information can pass from one state to another. The map has no information about adjacency. The organism can navigate by learning to associate particular actions with particular states, but it cannot solve problems that require spatial extrapolation or insight. The other three maps show essentially complete linkages between states; these maps contain adjacency information and also some information about distance.

² Global Positioning System: a satellite-based system used in navigation and robot orientation.

³ For simplicity, we denote each unit (node) in the map by a single subscript, even though the map is two-dimensional.

tions 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.

Details of the computer algorithms we used are given in the Appendix.

Spatial Search

The model defines a route to be followed within the map. The route is defined by an expectation surface that is altered dynamically both by the passage of time and by the model organism's experience with reward and nonreward. In this section we describe how the behavior of this model matches qualitatively the behavior of searching animals in both open-field and constrained conditions (barriers, mazes).

Open-Field Foraging: Area-Restricted Search

If rats are allowed to find buried food at a particular spot in a large enclosure, removed, and then returned after a delay, they usually begin digging close to the rewarded spot (cf. Cheng, 1992). If they fail to find food at once, the search is restricted to the immediately surrounding area before a more general, unlocalized search begins. Area-restricted search has been observed in a variety of species, including common shrews (Pierce, 1987), pigeons (Cheng, 1989), and even insect larvae and ants (Banks, 1954; Bond, 1980; Harkness & Maroudas, 1985; Nakamura, 1982; Wehner & Srinivasan, 1981).

The two-dimensional diffusion model produces area-restricted movement patterns when the simulated organism is reintroduced to the enclosure after finding reward at a particular spot. We begin our simulation with the assumption of a low, uniform level of reward expectation at every unit in the map. This small a priori expectation causes the organism to move around the enclosure in an exploratory way as successive failures to find food "zero" expectation at each spot. Once the model organism finds reward at a particular spot, expectation for the active unit is set to unity and the organism is immediately removed from the enclosure (i.e., all units made inactive). Reward expectation at that position immediately begins to diffuse according to Equation 2 at a rate determined by α . Figure 6 shows the pattern of movement produced when the organism is subsequently reintroduced to the empty (no-food) enclosure.

The four panels show the pattern of movement generated using four different values of α . The movement patterns were smoothed using splines to eliminate the abrupt changes in direction that result from a coarse-grained discrete map. In all panels, diffusion occurred for 30 time steps while the organism was out of the enclosure. This process creates a three-dimensional "landscape" of reward expectation (not depicted), centered on the rewarded location. For each panel the organism was subsequently reintroduced at Position 1,1 and allowed to move for 60 time steps. As before, Equation 2 was iterated once for each step taken.

The hill-climbing action rule moves the organism directly to the peak of the expectation landscape. This movement has the appearance of goal direction (i.e., the model organism moves straight to the rewarded spot), but in fact only local processes

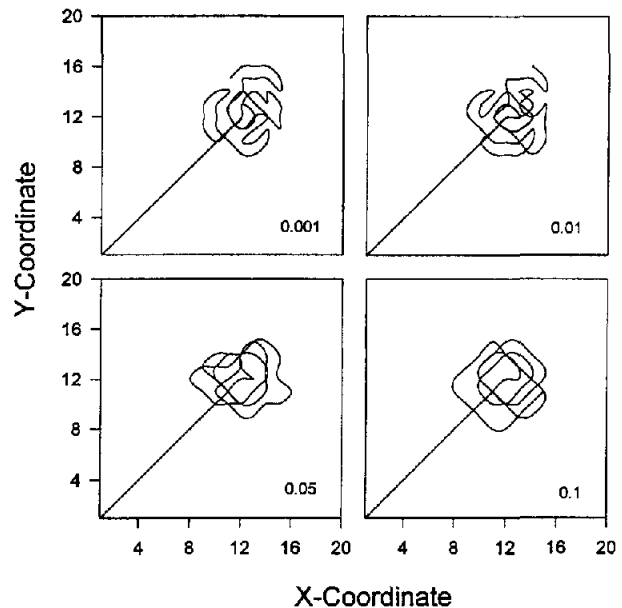


Figure 6. Area-restricted movement patterns. After "finding" food at location (12, 12; each quadrant), the simulated organism was removed from the enclosure. The figure shows the movement patterns produced when the organism is subsequently reintroduced to the empty (no-food) enclosure at position (1, 1; each quadrant) and allowed to move for 60 time steps. The different panels represent movement patterns produced with different values of the diffusion-rate parameter, α , shown in the bottom-right corner of each quadrant.

are involved: The organism cannot sense the peak of the expectation landscape from a distance. At each step, the value of the reward expectation for the active unit is set to zero, because no reward is obtained. The diffusion process allows reward expectation to "flow" back gradually into the units the organism has already visited (like tracks made by walking through a puddle of viscous fluid). The amount of flow back into previously visited spots depends on the amount of reward 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 of the expectation landscape act like a working memory that shows spontaneous recovery.

With each unrewarded step, reward expectation is zeroed at that point. As more and more steps are made near the rewarded spot, the overall level of reward expectation in the area declines, and the organism 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; see Figure 7).

The area-restricted "search" pattern of movement is centered on the rewarded 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. This general pattern is not dependent on the value of the diffusion-rate parameter α .

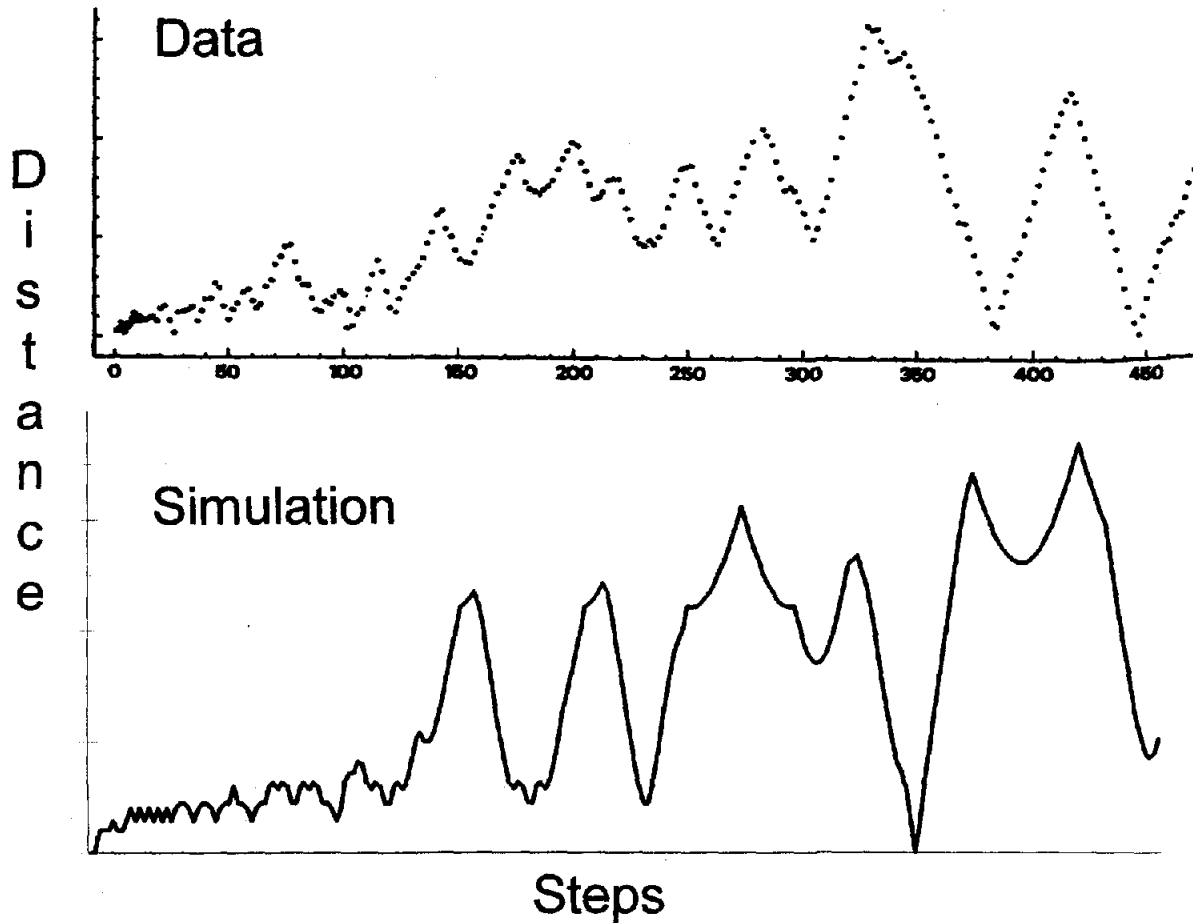


Figure 7. Top panel: Distance from the origin ("nest") of a single searching ant at 2-m intervals as a function of total path length, in m (from Wehner & Srinivasan, 1981, Figure 8b). Bottom panel: Distance from goal as a function of time steps of a simulated organism searching according to the diffusion 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, $\alpha = 0.1$). We have found very similar patterns with α values ranging from 0.02 to 0.1 and initial iterations from 20 to 50.

Area-Restricted Search: Fine Structure

Desert ants of the genus *Cataglyphis* find their nest (a hole less than 1 cm in diameter, in the ground in an almost featureless environment) by dead-reckoning (also termed *path integration*). Unlike many other ant species, they do not use odor trails.⁴ Wehner and 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 of the sort predicted by our model.

A typical experiment was done 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' cognitive map is established through dead-reckoning (i.e., by integration of the path they have taken in leaving the nest, rather than via landmarks or other environmental features), the ant is completely unaffected by being displaced and 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 plots in Figure 6. An alternate strategy was to take a returning ant from the mouth of the nest: In this case, the animal begins search at once since its dead-reckoning system has been reset to zero by finding the nest.

⁴ The suggestion has been made that the ants may actually detect an odor emanating from the nest, and orient in this way, but we have read no suggestion that this is the case. It seems in any case unlikely because such a signal would be easily distorted by wind, which is common in desert habitats, and the resulting path alteration would be obvious to experimenters.

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 at different azimuthal directions. This system of loops is centered 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 origin⁵ . . . the return to the centre must be regarded as an intrinsic feature of the search strategy itself" (p. 319).

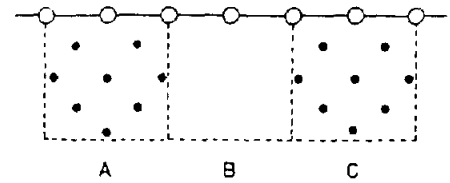
The best way to see this looping pattern is to plot the ant's distance from the origin ("nest") as a function of time (or, equivalently, path length). A typical example is shown in the top panel ("Data") of Figure 7, which is taken from Wehner and Srinivasan's (1981) Figure 8b. The figure shows distance from the origin on the y-axis versus path length (in m) on the x-axis. The division between "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 record of distance versus path length. 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 ("Simulation") is shown in the lower panel of Figure 7, and the general pattern is very similar: Focal search followed by wide return swings farther and farther away from the goal. We observed this same pattern with a range of α values and numbers of steps (see figure note). The reason the model shows periodicity like the data is well described by Wehner and Srinivasan (1981), who propose a stochastic optimal-search model with properties very similar to the diffusion model. They write:

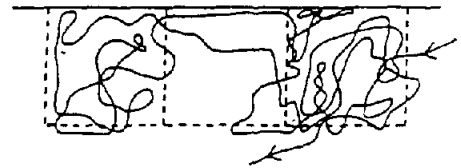
There can never be complete certainty . . . that a given region has been fully explored. All that can be said is that the *probability* of encountering unexplored areas within the region decreases as the time spent in searching the region increases. . . . Another way of expressing this idea is to say that the probability of finding the nest with [in] 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 (1981) propose an explicit stochastic optimal-search model that embodies this idea and it resembles their data pretty well (see their Figures 15, 16, 19, and 20), although not, to our eye, quite as well as the diffusion model (the oscillations of their model are too regular). But the diffusion model perfectly captures their 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 combi-

Procedure:



Search Path:



Simulation:

Search Path:

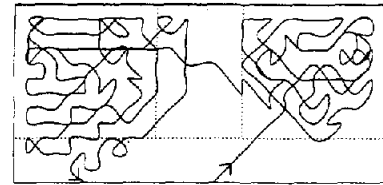


Figure 8. *Top panel:* Mellgren and Roper's (1986) procedure showing the placement of shelled peanuts in an open field. *Middle panel:* The path taken by a single European badger on the 6th night as it foraged for the peanuts. *Bottom panel:* The movement pattern produced by our model organism as it "foraged" for reward distributed in the same pattern as in the top panel. Dashed lines mark the three study areas.

nation that lead to periodical returns to the origin: At first, a sector (what Wehner and Srinivasan call an "azimuthal direction") 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 as the search moves outward.

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 6 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 8. The second panel shows the track of the

⁵Note that finding the nest resets the orientation mechanism; failing to find the nest at the origin does not.

badger on the 6th 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 8 shows a typical simulation result. We distributed the simulated rewards in the same spatial pattern as in the Mellgren and Roper (1986) experiment (top panel of Figure 8). As in the prior simulations, we allowed the organism to locate the rewards and immediately "removed" it from the area. In Mellgren and Roper's experiment, sessions were separated by approximately 24 hr, and the badger typically spent about 15 min in the study area collecting the peanuts. We simulated the 24-hr period by iterating the diffusion equation 300 times with a low diffusion rate ($\alpha = 0.001$) before returning the organism to the area.⁶ 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. Because 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 (1986) experiment, as well as our simulation, 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 shortcuts (i.e., to take a novel, shorter path after experience with a roundabout path to a goal). Chapuis (1987) did a very simple shortcut experiment. Imagine three open-field locations arranged as in each panel of Figure 9: A, B, and X, arranged in a triangle, with X (open circle) as starting point. A is closer to X than B: Distance $XA < XB$. A hungry dog on a leash was led from starting point (X) to Location A where food had been placed, and then led back to the starting location without being allowed to eat the food (Path XAX). This procedure was repeated for another, more distant, location, B (Path XBX). Path XAB is obviously shorter than the sum of the two training paths, XAXB (see Figure 9). However, going straight to B after eating the food at A involved taking a shortcut, Path AB, over ground previously unvisited by the dogs. Most dogs tested nevertheless took the shortcut when allowed to run free from the starting point.

The diffusion model works like this in the Chapuis (1987) 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 AB line. Figure 9 shows typical movement patterns produced by the model. The open circle in each panel is the starting point, X, and the two filled circles are the two reward locations, A

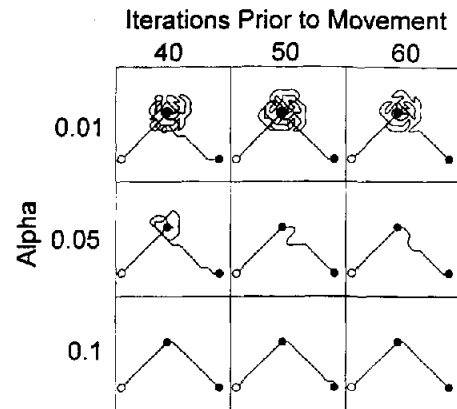


Figure 9. Simulations of Chapuis's (1987) shortcut experiment. The open circle in each panel is the starting point, X, and the two filled circles are the two reward locations, A and B. Rows show the effect of different values of the diffusion parameter, α ; columns show the effect of different numbers of iterations after seeing food at A and B before beginning to search again from Point X.

and B. Rows show the effect of varying the diffusion-rate parameter, α ; columns show the effect of varying the number of iterations after food at both locations has been "seen" but before the simulated organism begins foraging from the starting point, X. In every case, the model goes straight to A, the closer food source, first. Low values of α produce substantial area-restricted search around point A after the food is consumed, higher values produce little. In all cases, after searching much or little at A, the model goes straight to B. Thus, given an accurate map, the simulated organism, like Chapuis's dogs, will go straight to B rather than retracing its path to X. The model shows spatial "insight."

Menzel (1973) has done much more extensive shortcut, multiple-goal experiments with chimpanzees along these same lines, with results similar to Chapuis's (1987). We have not attempted to simulate these studies, because of the complicated cue environment and uncertainties about exact timing of events. But Menzel's results do not seem to be in any way incompatible with our analysis.

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. In the 1930s, the detour problem occasioned much debate between

⁶ Note that the basic diffusion model is not context sensitive and has no long-term memory. Without ad hoc 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 an ad hoc way of building in the idea that memory changes more slowly when the animal is outside the learning context than when it is inside. It is a crude way to implement the distinction between working and reference memory.

competing learning theories such as Hull's early S-R theory (Hull, 1938), Lewin's field theory (Lewin, 1933, 1935), and Tolman's "means-end-fields" (Tolman, 1932). (See, for example, Rashotte's excellent 1987 review.)

Simulation of open-field behavior assumes a fixed cognitive map. Learning about barriers obviously requires modification of the map. We have already described how our model incorporates barriers into the map. Encountering a barrier breaks the link between two adjacent units: parameter κ_{ij} is set equal to zero and N_i reduced by 1 if a barrier is encountered when the model attempts to move from node i to node j or the reverse. Reward expectation will not flow directly between the units, and no movement will be attempted between them. Note that we are not assuming what we hope to prove (that the model can solve Umweg problems), because encountering a barrier does not affect the model organism's behavior directly, it only changes the map. As we will see, there is an indirect effect, however, because of the effect of the changed map on patterns of diffusion and, therefore, on 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 (see Figure 10) between the subject and a reward. Because 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. Our model suggests that the difference between smart and less smart subjects lies 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.

Figure 10 shows simulations of the U-detour problem. We began the simulations 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. As we saw earlier, in a free field this initial condition produces an exploratory pattern of movement. As before, we allowed the diffusion equation to iterate 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 (κ_{ij} set to 0, N_i reduced by 1).

The simulation continued until the barrier had been encountered at each point. Subsequently, reward 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.

The panels in Figure 10 show four simulations with two starting positions (filled circles) and two goal positions (filled squares). In each case, the simulations produce a movement path that approximately minimizes the distance between the starting position and the reward location, given the constraints imposed by the lattice structure of our map. These movement patterns result from the diffusion of reward expectation around the barrier. The patterns do not depend on the value of the diffusion rate parameter, α . The movement path closely resembles reports in the Hull, Lewin, and Tolman studies cited earlier, 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 blind⁷ 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 two panels of Figure 11 depict Dennis's apparatus, as well as tracings from a typical rat during five consecutive trials with one or two partitions in place. The start box (not shown) was located at the upper left corner, and the goal box was at the bottom right corner.

The two panels of Figure 12 show the movement patterns produced by our simulation of Dennis's (1929) procedure. Once again, we began the simulations with the assumption that all adjacent units are 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

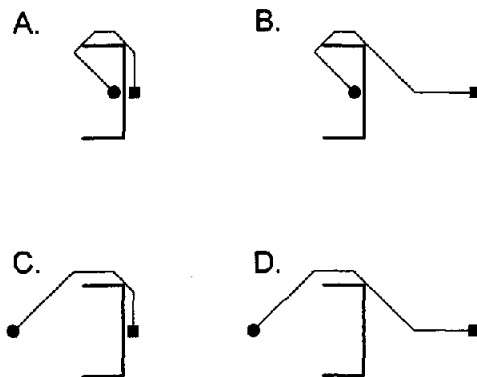


Figure 10. Barrier problem. A U-shaped barrier lies between the simulated organism (filled circle) and the reward (filled square). The four panels show two starting positions (Panels A and C) and two goal positions (Panels B and D). The organism was allowed to encounter the barrier and find reward; it was then removed from the apparatus and returned to the starting position. Each panel shows the path to the goal taken by the simulated organism when returned to the start.

⁷ 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.

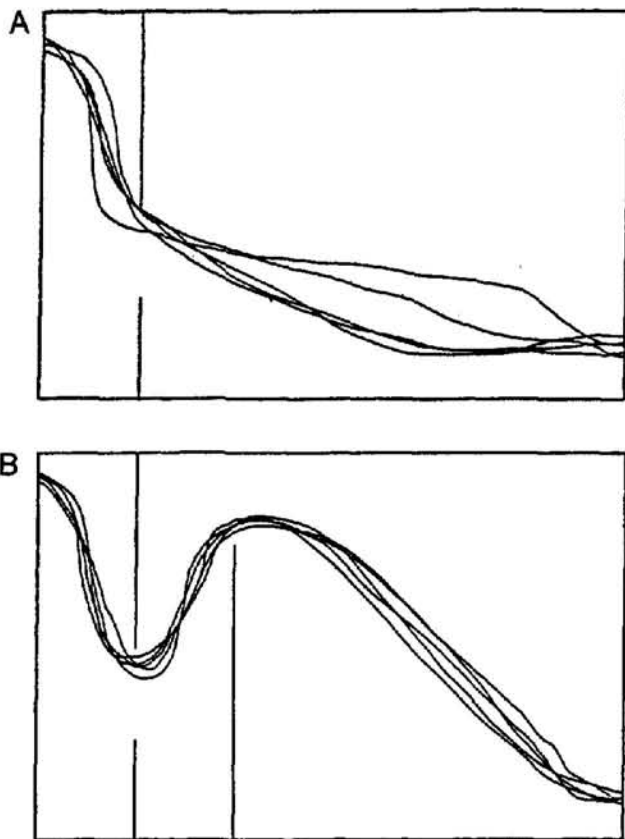


Figure 11. Dennis (1929) traced the paths taken by a blind rat in five consecutive trials in an apparatus containing one (Panel A) or two (Panel B) removable partitions. The start box (not shown) was located at the upper left corner, and the goal box was at the lower right corner.

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. We allowed the simulation to iterate until all the barriers had been identified. Subsequently, we removed the simulated rat from the enclosure and placed it in the start box. When the rat's continued exploration eventually led it to the goal box, it was immediately rewarded 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 rat explored the enclosure. Movement was mostly restricted to the walls and partitions. The organism 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 (Hall & Ballachey, 1932; Patrick & Laughlin, 1934).

After the simulated organism 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 (cf. top panel of Figure

11), the simulation produced movement that approximately minimized the number of steps from the start box to the goal box. The model achieved a similar minimum-distance solution when a second partition was added (bottom panel). The main difference between the movement produced by our simulation and that of actual rats is the occasional abrupt changes in path direction because of the coarse grain of our map: The finer the map, the closer is the predicted movement pattern to the minimum-distance solution.

After Dennis's (1929) rats had been trained to run through the enclosure with one partition in place, the partition was removed and the trials continued. The top panel of Figure 13 shows the patterns produced by the first run of Rat 1, the second run of Rat 2, the third run of Rat 3, etc. (Unfortunately, Dennis does not show data from individual subjects in consecutive trials, and it is not possible from his tracings to be sure about the order in which the paths crossed the location where the partition had been.) Dennis described the results this way:

The nature of the pathway followed when the cross-wall was lifted is shown by Figure [13]. The path which they had followed from 87 to 100 trials gave way immediately to a shorter one. 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

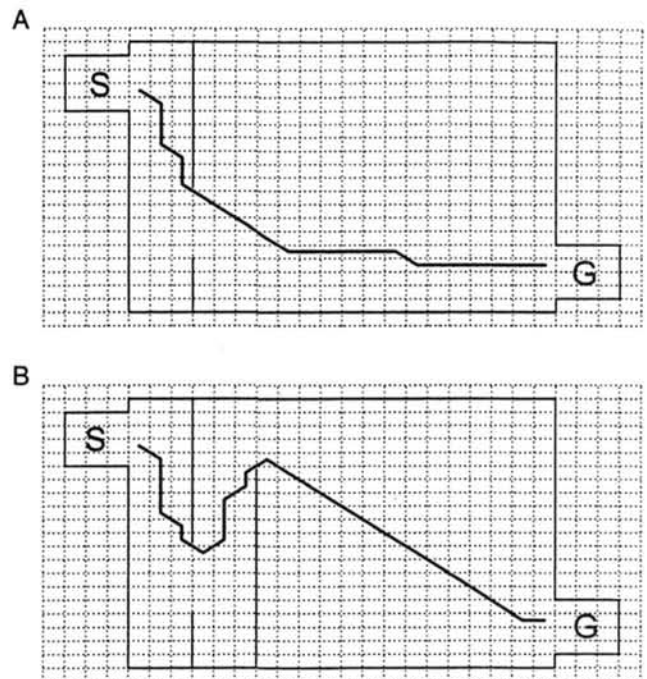


Figure 12. Movement patterns produced by our simulation of Dennis's (1929) procedure. See text for details. Panel A represents the movement of the simulated organism when one barrier is in place. Panel B represents the movement when two barriers are in place. The light dotted lines provide a visual representation of the units in the cognitive map. The simulated organism approximately minimized the distance between the start (S) and goal (G) boxes within the constraints of the cognitive map.

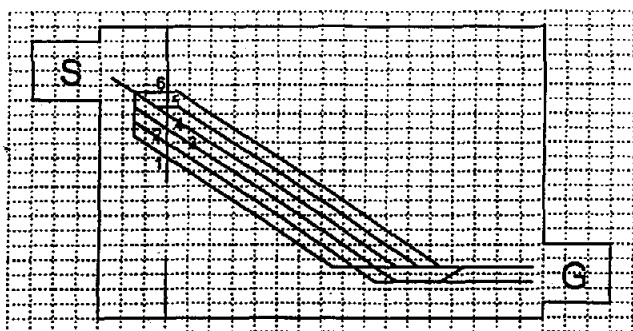
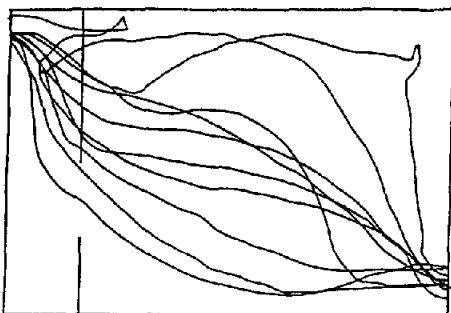


Figure 13. *Top panel:* Paths taken by Dennis's (1929) blind rats when the partition was removed. The patterns represent the first run of Rat 1, the second run of Rat 2, the third run of Rat 3, and so on. Unfortunately, Dennis did not show data from individual subjects in consecutive trials, and it is not possible to be sure about the order in which the paths crossed the location where the partition had been. *Bottom panel:* Movement patterns produced by our simulation of Dennis's (1929) procedure after the partition was removed. The six pathways represent consecutive rewarded trials. The order in which the paths occurred is labeled. In each trial the distance traveled from the start box (S) to the goal box (G) was shorter than when the partition had been in place, in general agreement with Dennis's data.

toward a straight line, but the behavior varied from trial to trial and from rat to rat. (p. 70)

Dennis's (1929) description resembles several features of the simulation in the bottom panel of Figure 13, which shows the paths produced when the partition is removed. The six labeled pathways represent consecutive rewarded trials. In each trial the distance traveled from the start box to the goal box was shorter than when the partition had been in place, in general agreement with Dennis's data. The simulations also show a similar effect of the prior "habit": Pathways were to the right of a straight line between the start and goal boxes. Finally, the simulation produced a similar progression toward a straight line. After the sixth trial, simulated paths generally overlapped Pathways 4-6.

Dennis (1929) 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. Our simulation results 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. We have already shown that our essentially local model is nevertheless capable of producing minimum-distance paths under many conditions. We show in this section three other examples of intelligent maze behavior that can be simulated by the diffusion model.

Tolman's "insight" mazes. The shortcut 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 14. 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. 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 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 dogleg in the original maze. Tolman et al. (1946) concluded that training produces (in some rats, at least) a "disposition to take the shortest Euclidean Path to the goal" (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 simulated organism 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 simulated organism will move

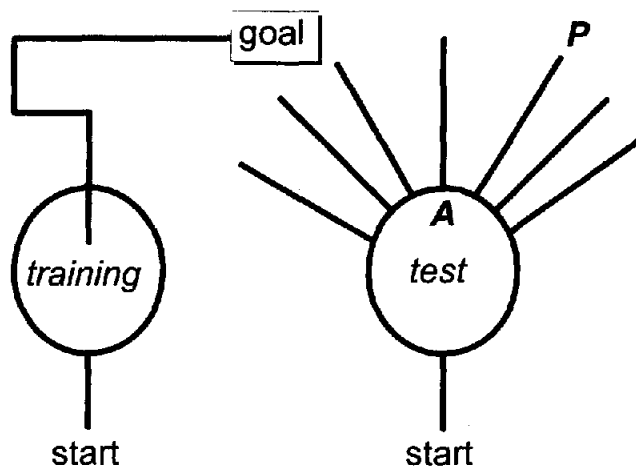


Figure 14. *Left:* Schematic picture of the training maze for rats used by Tolman, Ritchie, and Kalish (1946). *Right:* The "starburst" maze used by Tolman et al. (1946) during testing with a block at A. P = the path pointing directly at what used to be the goal.

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 system to explore again the edges of the open field, where there 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 Tolman and Honzik (1930) "insight" problem with which we began (Figure 1) can obviously be solved in the same way. Given an accurate map and knowledge of a barrier at Point B, expectation can 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 they revisit no arm until all have been explored. This behavior is sometimes termed *paradoxical* because 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, 1983, 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, 1932, 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, as we have also seen, yields exploratory behavior—and would yield spon-

taneous 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 (cf. Figures 6 and 8), 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. Hence, the model, like the real animal, will hesitate before searching again once all arms have been visited on a given day.

Thus, the 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 (1932) 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 (expectation surfaces) at the beginning of testing with the two maze types; it need not reflect some special set of processes unique to rats.

Discussion

Once an organism has the capacity to form some kind of internal, cognitive map, a very simple, behavioristic 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.) The route finder we propose is nothing but a dynamic formulation of stimulus generalization, a basic behavioristic process. Thus, Tolman's (1932) contention that something more than S-R principles is required to account for maze behavior is only partly true. A map is required (in our scheme) but beyond that, no "insight," no internal Cyclopean eye with an overview of the situation-as-a-whole, is necessary.

Moreover, the idea of an internal representation is not in any way foreign to behavioristic ideas. Stimulus generalization assumes representation of a stimulus dimension, and evidence abounds that even a psychological property as apparently primitive as color is not captured by a simple physical dimension, such as wavelength. Generalization takes place not in physical but in psychological space (cf. Shepard, 1965). Physical space is internally represented, even by the lowly ant. But once this is conceded, little more need be assumed to explain apparently complex spatial behavior.

There are several omissions in our account. Most glaring is the absence of a system for the formation of a map. Yet this may not be a very formidable problem. If we restrict ourselves to landmark learning, for example, then all that is necessary is a process that maps the space of landmark azimuths (two-dimensional for two landmarks, three-dimensional for three, etc.) on to a two-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. In a related article, Schmajuk and Thieme (1992) have proposed a neural-network route finder that uses recurrent links between "views" to propagate binary associations across the map, a process that seems to have some similarities to diffusion in our model and also suggests a way to combine map reading with map formation.

Hampson (1990) has proposed a connectionist scheme that also combines map building and route finding. Instead of diffusion, Hampson used forward and backward "spreading activation" to create "subgoals," which guide movement. Like our model, Hampson's broke links between nodes separated by a barrier. Unlike ours, it also created the map, by making links between nodes representing places visited through random movement: It built the map, as well as constraining it (as we do by breaking links after experiencing barriers). (Nevertheless, for most of his simulations, "to simplify matters and avoid the initial learning phase," Hampson provided "a complete representation of the spatial state space" (p. 151). The most important difference between our model and the models of Hampson and Schmajuk et al. is that our model organism learns something about the presence or absence of reinforcement *with each step*—reinforcement sets expectation of the active unit to a positive value, but nonreinforcement sets it to zero. This process gives rise to a dynamic expectation landscape that changes with every step the organism takes. The models of Hampson and Schmajuk et al. create gradients the model organism will climb, but movement does not affect gradient shape. Once created, the expectation gradients are climbed, but not directly altered, by the organism's movement until the peak is reached, at which point the simulation typically stops. As a result, these models are likely to make similar predictions to ours in mazes with runways that are sufficiently narrow (i.e., the Tolman and Honzik maze) to preclude an effect of diffusion from the organism's unreinforced steps. But predictions will diverge in mazes with more spacious areas (cf. the starburst maze and Dennis's mazes), and especially in the open field. In the open field, these models are unlikely to show the path self-crossings and the patterns of area-restricted search characteristic of the diffusion model—and the data (cf. Figures 6–9). Because of their static expectation surfaces, the Hampson and Schmajuk et al. models may also be subject to trapping in local maxima, something that cannot occur with our model. Unfortunately, Hampson did not present simulated search paths that can be compared directly with data and with our predictions. Hampson's theory also fails in some interesting ways. For example, he wrote in connection with open-field search, "The predator will follow a gradient that is determined by a single "best" goal. The result is that it will reject a cluster of goals in favor of a single, slightly stronger goal at the same distance" (p. 160). Apparently he could find no easy fix for this problem. Because it is basically a spatial generalization model, the diffusion model does not make this particular error. It will combine the expectancies of adjacent small goals so that they can win out over a single large goal at the same distance.

Another omission in our model is a comprehensive way of dealing with temporal properties of reinforcement. For example,

Hearst, Koresko, and Poppen (1964) showed that generalization gradients measured after training with very intermittent reinforcement (e.g., a variable-interval, 4-min schedule) are shallower than gradients obtained after less intermittent reinforcement (e.g., VI30 s). The diffusion model generates gradients with lower peaks after less frequent reinforcement, but the normalized gradients are not shallower. A diffusion process that incorporates a property we have termed *rate sensitivity* (Staddon & Higa, 1996; Staddon, 1998) may be necessary to deal with these effects. Rate sensitivity refers to a well-known effect of intermittent reinforcement, that although it produces less responding under training conditions, it may produce more persistent responding when reinforcement is withheld (the partial reinforcement-extinction effect). Simple diffusion is not rate sensitive in this sense.

Although our 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. 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, the model cannot accommodate histories that are likely to have very different effects on future behavior. Real animals are likely to 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 seems to be an example of a more general problem that Grossberg (1987) has termed the *stability–plasticity dilemma*: the stability of learning in the face of new information. One way to deal with it is to add another layer of units whose job is to learn about the histories of units in the first layer. Carpenter and Grossberg (e.g., 1988) have made a well-known proposal along these lines, and doubtless others are possible.

Our model is sensitive to only part of the stimulus context. We assume that in the absence of its stimulus a unit is inactive, but it is also likely that diffusion rate will vary as a function 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) The model is incomplete in the ways just discussed—it lacks context sensitivity as well as a map-making process. (b) Even though the model is entirely deterministic, nonlinearity in the reinforcement assumptions means that the predicted path is exquisitely 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 a principled way of setting 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 (more on this in a moment).

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's (1987) two-goal experiment, for example (Figure 9), the model should be able to predict search paths in similar experiments with more or differently arranged goals. Our assumptions about barrier learning could be tested systematically by varying exposure to barriers and looking at detour paths (cf. Figure 10).

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 (1987) study the dogs went to the closer food location before they went to the farther location. However, our 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. The following is another example: If the experimenter shows the dog the more distant food, waits a little while, shows the dog a second, closer food, and then releases the dog without delay at the starting location, the model implies that diffusion from the second location may not have had time to reach the starting location, so that the animal goes to the farther location first.

The following is a final example: In the open field, gradients from adjacent goals may combine. So if two food locations are relatively close to each other and there is very little delay before releasing the animal from the starting location, it should choose the closer food source. But with increased delays, the two gradients will combine into one large hill, with a peak near the midpoint between the two locations. Therefore, the animal should go initially to the midpoint before changing direction.

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 because 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.

Despite these limitations, the simplicity of the diffusion model and its ability to generate complex search paths whose general properties closely resemble those of real animals in a variety of environments recommend it as a useful component for any comprehensive model of spatial behavior.

References

- Banks, C. L. (1954). The searching behaviour of coccinellid larvae. *Animal Behaviour*, 19, 58–86.
- Banks, R. B. (1994). *Growth and diffusion phenomena: Mathematical frameworks and applications*. New York: Springer.
- Berg, H. C. (1983). *Random walks in biology*. Princeton, NJ: Princeton University Press.
- Bond, A. B. (1980). Optimal foraging in a uniform habitat: The search mechanism of the green lacewing. *Animal Behaviour*, 28, 10–19.
- Carpenter, G. A., & Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, March, pp. 77–88.
- Chapuis, N. (1987). Detour and shortcut abilities in several species of mammals. In P. Ellen & C. Thinus-Blanc (Eds.), *Cognitive processes and spatial orientation in animal and man* (pp. 39–54). Boston: Nijhoff.
- Cheng, K. (1989). The vector sum model of pigeon landmark use. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 366–375.
- Cheng, K. (1992). Three psychophysical principles in the processing of spatial and temporal information. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 69–88). Hillsdale, NJ: Erlbaum.
- Dale, R. H. I. (1982). Parallel-arm maze performance of sighted and blind rats: Memory and maze structure. *Behaviour Analysis Letters*, 2, 127–139.
- Dashiell, J. F. (1930). Direction orientation in maze running by the white rat. *Comparative Psychology Monographs*, 7 (No. 2).
- Dawkins, M. (1982). Evolutionary ecology of thinking: State of the art report. In D. R. Griffin (Ed.), *Animal mind—human mind*. Berlin: Springer.
- Dennis, W. (1929). The sensory control of the white rat in the maze habit. *Journal of Genetic Psychology*, 36, 59–90.
- Deutsch, J. A. (1960). *The structural basis of behavior*. Chicago: University of Chicago Press.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT/Bradford.
- Gould, J. L., & Gould, C. G. (1982). The insect mind: Physics or metaphysics? In D. R. Griffin (Ed.), *Animal mind—human mind* (pp. 269–297). Berlin: Springer.
- Grossberg, S. (1987). *The adaptive brain*. (2 vols.). Amsterdam: Elsevier.
- Guthrie, E. R. (1935). *The psychology of learning*. New York: Harper.
- Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, 51, 79–88.
- Hall, C., & Ballachey, E. L. (1932). A study of the rat's behavior in a field: A contribution to method in comparative psychology. *University of California Publications in Psychology*, 6(1), 1–12.
- Hampson, S. E. (1990). *Connectionist problem solving: Computational aspects of biological learning*. Boston: Birkhäuser.
- Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, 58, 321–333.
- Hanson, H. M. (1961). Stimulus generalization following three-stimulus discrimination training. *Journal of Comparative and Physiological Psychology*, 54, 181–185.
- Harkness, R. D., & Maroudas, N. G. (1985). Central place foraging by an ant (*Cataglyphis bicolor* Fab.): A model of searching. *Animal Behaviour*, 33, 916–928.
- Hearst, E., Koresko, M. B., & Poppen, R. (1964). Stimulus generalization and the response–reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 7, 369–380.
- Hoffman, H. S., & Fleshler, M. (1961). Stimulus factors in aversive controls: The generalization of conditioned suppression. *Journal of the Experimental Analysis of Behavior*, 4, 371–381.
- Hull, C. L. (1938). The goal-gradient hypothesis applied to some "field-force" problems in the behavior of young children. *Psychological Review*, 45, 271–299.
- Kalish, H. I., & Guttman, N. (1959). Stimulus generalization after training on three stimuli: A test of the summation hypothesis. *Journal of Experimental Psychology*, 57, 268–272.

- Lett, B. T. (1975). Long-delay learning in the T-maze. *Learning and Motivation*, 6, 80–90.
- Lewin, K. (1933). Vectors, cognitive processes, and Mr. Tolman's criticism. *Journal of General Psychology*, 8, 318–345.
- Lewin, K. (1935). *A dynamic theory of personality: Selected papers*. New York: McGraw-Hill.
- Mellgren, R. L., and Roper, T. J. (1986). Spatial learning and discrimination of food patches in the European badger (*Meles meles L.*). *Animal Behavior*, 34, 1129–1134.
- Menzel, E. (1973). Chimpanzee spatial memory organization. *Science*, 182, 943–945.
- Morris, R. G. M. (1981). Spatial location does not require the presence of local cues. *Learning and Motivation*, 12, 239–260.
- Nakamura, K. (1982). Switchover in searching behaviour of *Coccinella septempunctata L.* (Coleoptera: Coccinellidae) caused by prey consumption. *Applied Entomology and Zoology*, 17, 501–506.
- Olton, D., & Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 2, 97–116.
- Olton, D., Walker, J. A., Gage, F. H., & Johnson, C. (1977). Choice behavior of rats searching for food. *Learning and Motivation*, 8, 315–331.
- Patrick, J. R., & Laughlin, R. M. (1934). Is the wall-seeking tendency in the white rat an instinct? *Journal of Genetic Psychology*, 44, 378–389.
- Pavlov, I. P. (1927). *Conditioned reflexes*. New York: Oxford University Press.
- Pierce, G. J. (1987). Search paths of foraging common shrews *Sorex araneus*. *Animal Behavior*, 35, 1215–1224.
- Rashotte, M. E. (1987). Behavior in relation to objects in space: Some historical perspectives. In P. Ellen & C. Thinus-Blanc (Eds.), *Cognitive processes and spatial orientation in animal and man* (pp. 39–54). Boston: Nijhoff.
- Reynolds, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, 4, 57–71.
- Riley, D. A. (1968). *Discrimination learning*. Boston: Allyn and Bacon.
- Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103, 736–760.
- Schmajuk, N. A., & Blair, H. T. (1993). Place learning and the dynamics of spatial navigation: A neural network approach. *Adaptive Behavior*, 1, 353–385.
- Schmajuk, N. A., & Thieme, A. D. (1992). Purposive behavior and cognitive mapping: A neural network model. *Biological Cybernetics*, 67, 165–174.
- Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, 65, 242–256.
- Shepard, R. N. (1965). Approximation to uniform gradients of generalization by monotone transformations of scale. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 94–110). Stanford, CA: Stanford University Press.
- Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 44, 430–444.
- Staddon, J. E. R. (1977). Behavioral competition in conditioning situations: Notes toward a theory of generalization and inhibition. In H. Davis & H. M. B. Hurwitz (Eds.), *Operant-Pavlovian interactions* (pp. 103–131). Hillsdale, NJ: Erlbaum.
- Staddon, J. E. R. (1983). *Adaptive behavior and learning*. New York: Cambridge University Press.
- Staddon, J. E. R. (1985). Inference, memory and representation. In N. M. Weinberger, J. L. McGaugh, & G. Lynch (Eds.), *Memory systems of the brain: Animal and human cognitive processes* (pp. 287–295). New York: Guilford Press.
- Staddon, J. E. R. (1998). The dynamics of memory in animal learning. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in Psychological Science: Vol. 2. Proceedings of the XXVI International Congress of Psychology*. Montreal, Quebec, Canada (pp. 259–274). Psychology Press, Hove, Sussex, England.
- Staddon, J. E. R., & Higa, J. J. (1991). Temporal learning. In G. Bower (Eds.), *The Psychology of Learning and Motivation* (Vol. 27, pp. 265–294). New York: Academic Press.
- Staddon, J. E. R., & Higa, J. J. (1996). Multiple time scales in simple habituation. *Psychological Review*, 103, 720–733.
- Staddon, J. E. R., & Reid, A. K. (1990). On the dynamics of generalization. *Psychological Review*, 97, 576–578.
- Terrace, H. S. (1964). Wavelength generalization after discrimination learning with and without errors. *Science*, 144, 78–80.
- Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning* (pp. 237–275). Hillsdale, NJ: Erlbaum.
- Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.
- Tolman, E. C., & Honzik, C. H. (1930). "Insight" in rats. *University of California Publications in Psychology*, 4, 215–232.
- Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning: I. Orientation and the short-cut. *Journal of Experimental Psychology*, 36, 13–24.
- von Frisch, K. (1967). *Dance language and the orientation of bees*. Cambridge, MA: Harvard University Press.
- Wehner, R., & Srinivasan, M. V. (1981). Search behaviour of desert ants, Genus *Cataglyphis* (Formicidae, Hymenoptera). *Journal of Comparative Physiology*, 142, 315–338.

Appendix

Details of the Computer Simulations

Diffusion is a parallel process, so that implementation by serial computer always entails some compromises. The one-dimensional simulations (cf. Figures 2-4) were straightforward implementations of Equation 1, carried out on a spreadsheet. The two-dimensional simulations required more extensive programming. Physical locations were represented by individual nodes, and each node may have links connecting it to as many as eight neighboring nodes (cf. Figure 5). Every node has a value of reward expectancy associated with it. The values of all nodes in the map may change with every iteration of the diffusion equation. With an incomplete map (cf. Figure 5), some nodes may be unreachable from others—the map may consist of separate, unconnected graphs. Therefore, the diffusion process had to treat each graph as separate. Formally, each area was represented as an undirected graph (such as a tree in which the node representing the root has not been

identified). Diffusion should occur in all nodes of all graphs with every iteration. We met this requirement by implementing the task as a spanning forest, which is an algorithm that traverses all connected or unconnected graphs, visiting each node exactly once. It incorporates breadth-first traversal of each independent graph, and all links emanating from every node are considered. This algorithm allows each node's expectancy value to be updated by the diffusion equation exactly once per iteration, no matter how incomplete or complex the topological map may become.

Received July 14, 1995
 Revision received March 10, 1997
 Accepted March 10, 1998 ■



**AMERICAN PSYCHOLOGICAL ASSOCIATION
 SUBSCRIPTION CLAIMS INFORMATION**

Today's Date: _____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION _____			MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL) _____		
ADDRESS _____			DATE YOUR ORDER WAS MAILED (OR PHONED) _____		
CITY _____ STATE/COUNTRY _____ ZIP _____			<input type="checkbox"/> PREPAID <input type="checkbox"/> CHECK <input type="checkbox"/> CHARGE CHECK/CARD CLEARED DATE: _____		
YOUR NAME AND PHONE NUMBER _____			(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.) ISSUES: <input type="checkbox"/> MISSING <input type="checkbox"/> DAMAGED		

TITLE	VOLUME OR YEAR	NUMBER OR MONTH
_____	_____	_____
_____	_____	_____
_____	_____	_____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.

(TO BE FILLED OUT BY APA STAFF)

DATE RECEIVED: _____	DATE OF ACTION: _____
ACTION TAKEN: _____	INV. NO. & DATE: _____
STAFF NAME: _____	LABEL NO. & DATE: _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.