

Chapter 6: Optimal Foraging Theory: Constraints and Cognitive Processes

Barry Sinervo © 1997-2006

OPTIMAL FORAGING THEORY

THE PREY SIZE-THRESHOLD: A DECISION RULE THAT MAXIMIZES PROFIT
AN OPTIMAL DECISION RULE FOR CROWS FORAGING ON CLAMS
SIDE BOX 6.1. THE MECHANICS OF CONSTRAINTS ON OPTIMAL FORAGING
A SUMMARY OF THE MODEL BUILDING PROCESS
VARIATION IN FEEDING MECHANISMS WITHIN A POPULATION
INDIVIDUAL FORAGING CAN DETERMINED BY GENES, ENVIRONMENT, OR CULTURE
THE MARGINAL VALUE THEOREM AND OPTIMAL FORAGING
SIDE BOX 6.2: MARGINAL VALUE THEOREM: TRAVEL TIME AND ENERGY GAIN
CURRENCY AND LOAD SIZE IN PARENTAL STARLINGS
CURRENCY OF BEE COLONIES

FORAGING IN THE FACE OF A RISKY REWARD

FORAGING WITH A VARIABLE REWARD AND RISK-AVERSIVE BEHAVIOR
FOOD LIMITATION AND RISK-AVERSE VERSUS RISK-PRONE BEHAVIOR IN SHREWS
ADAPTIVE VALUE OF RISK-SENSITIVITY AND THE THREAT OF STARVATION.....
ARE ANIMALS RISK-SENSITIVE IN NATURE?
DECISION RULES AND COGNITION DURING FORAGING
COGNITION AND PERCEPTUAL CONSTRAINTS ON FORAGING?
RISK AVERSION AND THE COGNITIVE BUMBLE BEE
SIDE BOX 6.3. MEMORY LIST LENGTHS IN BUMBLE BEES.....
THE SEARCH IMAGE
PERCEPTUAL CONSTRAINTS AND CRYPTESIS

SUMMARY: OPTIMAL FORAGING AND ADAPTATIONAL HYPOTHESES.....
REFERENCES FOR CHAPTER 6

Optimal foraging theory

Chances are, when observing animals in the wild, you are most likely to see them foraging for food. If successful, their foraging efforts culminate in feeding. Animals search, sense, detect and feed. For humans, feeding is often associated with pleasure. Similar sensations may underlie the proximate drive that motivates feeding behavior of animals. However, the ultimate reason for feeding arises from the difference between life and death. At some point in an animal's life it may experience episodes of starvation and prolonged starvation can lead to death. If animals survive and die as a function of variation in their foraging strategies then natural selection has run its course. Animals that survive are able to contribute genes to the next generation, while the genes from animals that die are eliminated, and along with it unsuccessful foraging behaviors. Understanding the rules that shape the foraging behavior of animals has been a central focus of behavioral analysis for more than four decades (Pyke et al. 1977).

It seems reasonable to assume energy gain per unit of time might maximize the resource an animal has for survival and successful reproduction. For example, the common shrew, *Sorex araneus*, faces foraging decisions that keep it only a few hours away from death. Like all mammals, the shrew maintains a high and constant body temperature during activity. To keep itself warm, the shrew has a very active metabolism. Because the shrew's body is extremely small compared to larger mammals its surface area is very large relative to its body mass. If we compare the shrew to a rat, we would see that the shrew loses considerable heat to the environment. On a gram-per-gram basis the shrew's metabolism is much higher than mammals of larger size (Peters 1983; Calder 1984). Because of its high metabolism, the shrew satiates its voracious appetite with protein-rich insects, a high quality resource. The shrew must forage constantly, and barely has a moment to sleep because its small body size does not afford it the luxury of a thick layer of fat. The shrew has very few energy reserves onboard and only few hours without feeding (Barnard and Hurst 1987) can lead to death (Crowcroft 1957; Vogel 1976).

Even if food is abundant in the environment, and the shrew does not face life-or-death foraging decisions, it must have sufficient energy to reproduce. Natural selection will favor those individuals in a population

that have relatively high reproductive output. Thus, survival and reproduction must also be related to the efficiency of energy acquisition and energy storage. A reproductive female shrew has the added energetic outlay of nursing young. Reproductive females must maintain a positive energy balance for themselves and acquire enough excess energy to nurse their pups with energy-rich milk. The efficiency of the female shrew's foraging decisions may affect the size of her pups at weaning. In turn, the size at weaning might impact the probability of their survival to maturity. Life, death, birth, and successful reproduction in the shrew are measured in terms of calories taken in on a minute-by-minute basis.

Given the urgency of the “decisions” faced by shrews, the shrew may not even consider every insect it encounters as a worthwhile prey item. Imagine that a shrew is foraging for prey. During its forays in the rich humus of the forest floor, it encounters some small, but evenly dispersed species of grub with clockwork regularity. When it encounters one grub, should it eat the isolated prey? It takes some time to handle the prey and then more time to search for another. To calculate the value of that isolated prey item for the shrew we should take into account the value of the individual prey and the time it takes to find the prey. Should the shrew ignore the single isolated prey item or continue searching for a concentrated nest of termites that yields a much higher payoff. While the payoff from a termite nest is high, the nests are dispersed in the environment and locating them is a stroke of luck. The payoff from a large concentration of termites means the difference between making it through the long cold night versus the sure death it faces from feeding on the small grubs that are evenly distributed which it encounters on a regular basis. Yet these grubs are insufficient to sustain its needs. Animals make foraging decisions in the face of uncertainty. In this chapter, we address issues of foraging in the face of uncertainty. In other words, when does it pay to gamble? To understand gambling, we first need to understand the currency used by animals to make decisions, and the constraints on such decisions.

The theory of optimal foraging addresses the kinds of decisions faced by shrews, and indeed all animals. Regardless of whether foraging efficiency has an immediate impact on life or death, or whether it has a more cumulative or long-term effect on reproductive success, animals

make decisions in the face of constraints. **Temporal constraints** are couched in terms of the time it takes to find and process food. **Energetic constraints** are couched in terms of the metabolic cost of each foraging activity (foraging, processing, etc.) per unit time. Animals must learn about the distribution of food in their environment if they are to make the appropriate choices. How much learning is possible for an animal? Is there a limit to learning and memory, and do such **cognitive constraints** limit the foraging efficiency of animals?

The first issue we must address before considering the more complex decisions faced by economically minded, but perhaps cognitively challenged animals is the choice of **currency**. What are the units of currency used by animals when conducting their day-to-day transactions with the environment? How do basic energy and temporal constraints dictate the form of currency that animals use? A simple currency can be expressed in terms of the value of an item, taking into account the cost of acquiring the item, and the time taken to acquire the item. Natural selection might shape decision rules such that animals maximize net energy gain (e.g., gross gain - costs) as a function of time:

$$\text{Profitability of Prey} = \frac{\text{Energy per prey item} - \text{Costs to acquire prey}}{\text{Time taken to acquire prey item}}$$

The Prey Size-Threshold: a Decision Rule that Maximizes Profit

Prey size is one of the most conspicuous features that a predator could use to discriminate prey quality. The quality of the prey expressed in terms of energy content rises in direct proportion to mass, and corresponds roughly to the cubic power of prey length. It is more profitable to eat large prey, provided the prey is not too large so that the predator runs into **processing constraints**. For most animals the rule “never swallow anything larger than your head” is a simple rule by which to live. However, snakes break this rule routinely. Consider the anaconda in the Amazon forest that is capable of swallowing a deer. Some animals find ways around processing constraints by evolving adaptations. Snakes can eat things that are bigger than their head because they have a hinged jaw with an extra bone that gives them great flexibility when swallowing. All snakes share this unique adaptation for foraging. Most other animals solve the problem by chewing their food.

How about the rule, “eat things you can open.” The thickness of a shell may deter many predators. Most snakes can’t eat a bird egg because eggs are very resistant to radially distributed crushing forces (i.e., eggs must sustain the weight of the adult female during incubation). However, egg-eating snakes have evolved special points on the bottom of the spine (Arnold 1983). They press the egg up against the point and voila, cracked egg. A force that is concentrated at a point source breaks the egg like the edge of bowl used by humans. Egg-eating snakes have evolved an additional adaptation. Difficulty in opening or subduing prey should rise with prey size. Indeed, egg-eating snakes may have difficulty with an ostrich egg. The **handling time**, or the time taken to catch, subdue, and consume prey, will increase with prey size and prey armor.

If it is generally desirable to acquire large prey up to a **maximum size threshold**, the crucial question becomes what is **minimum size threshold** for prey in the diet. A prey item that is encountered in the environment should be consumed if it is above the size threshold, but should be rejected if it is below this threshold. The point at which consuming the prey becomes profitable depends on the **search time** and the **handling time** of prey as a function of the size threshold. If the size threshold is too large, a predator will wander around and deem a large fraction of the prey to be unacceptable. Such finicky behavior increases the search time between encountering successive prey items. This additional search time will eat away at the predators overall profits from a long sequence of prey, because the predator is metabolically active for a longer period of time during search. The predator receives no reward until it accepts and eats an item. The smallest size of prey that a predator should attempt to eat to maximize energy gain per unit time is our first example of an **optimal decision rule**, subject of course to the **constraints** of prey armor and the time taken to find prey.

An Optimal Decision Rule for Crows Foraging on Clams

The common crow, *Corvus caurinus*, forages in the intertidal and provides a clear example of optimal decision rules for prey-size selectivity (Richardson and Verbeek 1986). Japanese little-necked clams of various sizes are found under the sand on a typical ocean beach along the northwest coast of North America. The location of clams is not obvious to a crow and it has to probe to find clams. A crow spends an average of 34.6 seconds locating and digging up a single clam.

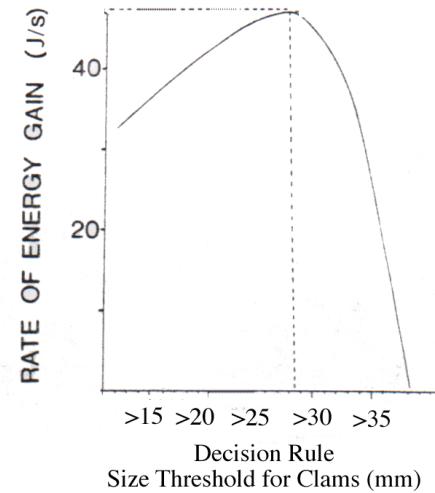


Figure 6.1. The rate of energy gain for crows, *Corvus caurinus*, foraging in the intertidal if they were to use different decision rules for the size threshold of clam that they are willing to open and eat. The greatest rate of energy gain or *optimal foraging strategy* would be achieved if crows ate every clam above 28.5 mm.

The crow has solved the problem of opening the clam with a short dive-bombing flight. The crow makes a short flight lasting 4.2 seconds and then drops the clam on a rock. If the clam does not break, the crow requires an additional 5.5 seconds for a second flight, and 2 more seconds for the second drop. It takes the crow an average of 1.7 flights to crack open a clam, thus the average clam requires $4.2 \times 1 + 5.5 \times 0.7 = 8.1$ seconds of flight time. The probability that a clam breaks open is independent of clam size. The amount of time that a crow invests in searching for clams is 4.3 times greater than the amount of time the crow spends in cracking open the clam with its dive-bomb flights. However the cost of flight in crows is nearly 4 times more expensive than the cost of search and digging. Thus, the search costs and handling costs expressed in terms of energy are nearly equivalent, but the search costs are more than 4 times more expensive than the handling costs expressed in terms of time. Crows reject many clams that they dig up and leave them on the beach unopened. If the crow goes to the trouble of finding and digging up a clam and all this takes time and energy, why doesn’t it eat all clams regardless of size, particularly since the search takes up the most time?

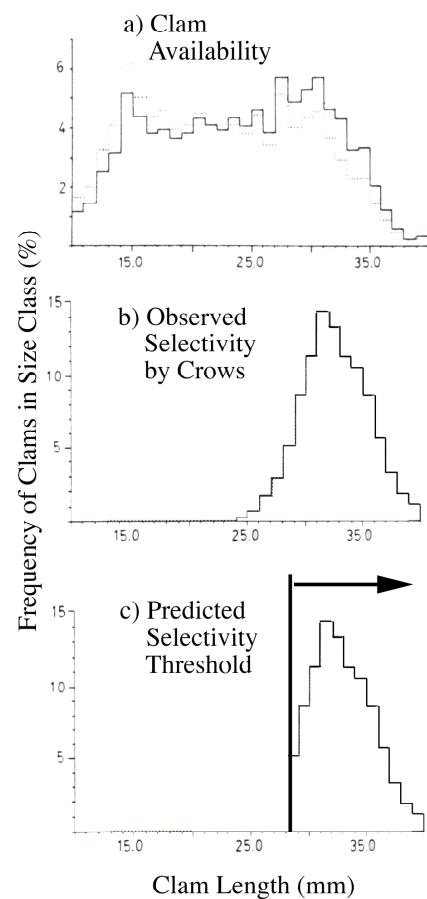
The answer to this question lies in the average net profitability of the clams as a function of size. We can compute the profitability of a single clam per unit of time once we discount all of the energy and time constraints of foraging by using the following equation:

$$\frac{\text{Energy}}{\text{Time}} = \frac{\text{Energy per clam as a function of size} - (\text{Search Costs} + \text{Handling Costs})}{(\text{Search Time} + \text{Handling Time})}$$

Figure. 6.2 a) Availability (frequency) of clams as a function of size on the beach at Mitlenatch Island, British Columbia. b) Frequency distribution of the clams that were eaten by the crows. c) Predicted size distribution of clams the crows should have chosen if they were foraging optimally and maximizing energy gain per unit time. The size-threshold is a truncation point or an absolute size below which the crows should not eat clams. Only a few clams were chosen below the size threshold and the majority of clams were chosen above the size threshold. The size-threshold is referred to as an *optimal decision rule*. From Richardson and Verbeek (1986)

To compute a clam's net profitability, Richardson and Verbeek (1985) computed the amount of energy that the crow expends in each of the following tasks: walking and searching, flying, and handling. This reflects the amount of energy expended in foraging. The amount of energy increases with the size of the clam and the net profitability of a single clam increases with size.

However, the simple formula in equation 6.2 is for the profits from a single clam. A crow eats many clams during a single bout of foraging, thus we must calculate the average profitability from a long string of rejected and accepted clams (Figure 6.1). The greatest rate of energy gain is achieved if a crow accepts clams greater than 28.5 mm. Why does energy gain decline when the crow uses a larger cutoff value for acceptable clams? Shouldn't such finicky behavior mean that it eats only the best and largest clams? Rather than show a formula, let's consider a verbal argument. A crow that is too choosy will wander across the



mudflat rejecting too many small clams. A crow that is not choosy enough will waste a lot of time feeding on tiny clams that take too much time to crack open for the measly reward found inside. If crows were to accept clams below this size threshold of 28.5 mm, they would take too long to open the clams relative to the energy content of derived from small clams. Below the optimal size threshold the *energy per clam* is so low that the return is not worth the handling time of flying over to the drop rock to crack the clam open. Conversely, rejecting too many large clams and using a decision rule above 28.5 mm would lead to more time spent searching for suitably large clams. Large clams constitute a much smaller proportion of the available clams, than medium sized clams. Increased search time lowers the average yield from all clams eaten.

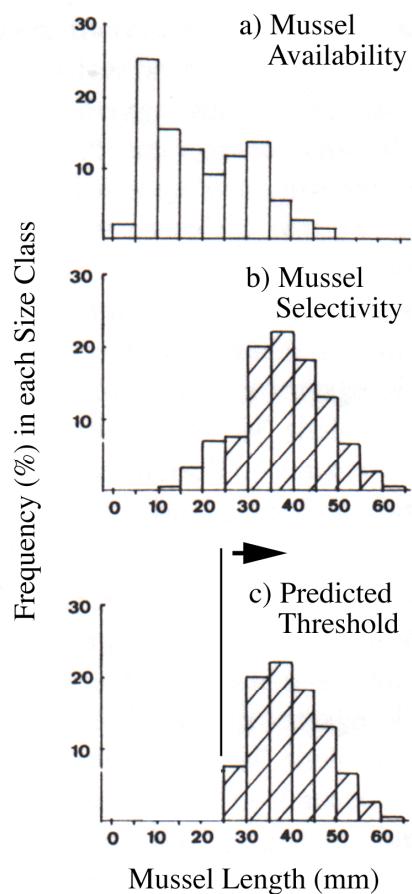
The best foraging strategy, or the **optimal decision rule** that crows should live by, is to accept all clams above 28.5 mm. It is always pays to attempt the largest clams because they are enormously profitable and do not require any extra energy to crack open. The size-threshold decision rule that was actually observed by Richardson and Verbeek (1985) was very close to 28.5 mm. How well does the model for the optimal foraging decision rule of clam selectivity fit the observed data? Only a few clams were chosen below this threshold, and nearly all clams were eaten above this size threshold. Crows appear to have an optimal decision rule for accepting and rejecting clams on the basis of size.

Oystercatchers and the Handling Constraints of Large Prey

Students of optimal foraging often seek generality by studying different species undertaking similar tasks. The foraging crows did not face any constraints of large prey size, however, the largest prey were relatively rare, forcing crows to feed on small clams to maximize profit. Meire and Ervynck (1986) carried out a similar analysis of Oystercatchers, *Haemotopus astralegus*, foraging in mussel beds as a test of optimal foraging theory. Oystercatchers forage on mussels with the added difficulty of cracking open the mussels with their bills rather than doing the fly-and-drop technique of crows. Oystercatchers appear to be quite size selective because large mussel have thicker shells. Even when they attempt to open the mussels with thin-walled shells, the oystercatchers have far lower success in opening large versus small mussel shells.

The increased handling time for the larger mussels enhances the relative profitability of small prey (see Side Box 6.1). If this difference in handling time were ignored and we used a model similar to the foraging crow, then oystercatchers should choose mussels greater than 55 mm in length. However, the enhanced profitability of the easy-to-open small mussels pushed the threshold value for the most profitability mussel down to a minimum size of 25 mm. Oystercatchers appear to use a decision rule that is very close to the size threshold predicted from an elaborate optimal foraging model that takes into account a number of constraints on foraging (see Side Box 6.1). Oystercatchers, like the crows, have developed an optimal rule for size selectivity feeding on mussels.

Figure. 6.3 a) Availability (frequency) of mussels, *Mytilus edulis*, as a function of size on Slikken van Vianen, a tidal flat in the Netherlands b) Frequency distribution of mussels that were observed as having been opened and eaten by oystercatchers (*Haematopus ostralegus*). c) Predicted size distribution of mussels the oystercatchers should have chosen if they were foraging optimally and maximizing energy gain per unit time. Only a few mussels were chosen below the optimal size-threshold and the majority of mussels were chosen above this point. See Side Box 6.1 for a complete description of constraints that Meire and Ervynck (1986) used in their optimal foraging model (from Meire and Ervynck, 1986).



Comparable tests of size selectivity have been repeated in a variety of taxa feeding on the same resource or drastically different resources. Shore crabs, *Carcinus maenas*, prefer to eat mussels of a size that maximize rate of energy return per unit of time (Elner and Hughes 1978). The theory of selectivity also appears to hold for herbivores that show selectivity for quality of plant food. Herbivores that range in size from the Moose, *Alces alces*, to the Columbian ground squirrel, *Spermophilus columbianus*, appear to be energy maximizers. However, the requirement for a balanced diet restricts herbivores from feeding exclusively on the highest energy foods, which lack vital micronutrients. Herbivores supplement their dietary energy gains with the right mix of alternative foliage that supplies key micronutrients (Belovsky 1978; Belovsky 1984). In contrast, predators can often follow a simple rule of eating prey that are made of things that they can use in building their bodies. Predators need not be as picky about the composition of prey, but as we have seen, can be quite sensitive to handling constraints.

A Summary of the Model Building Process

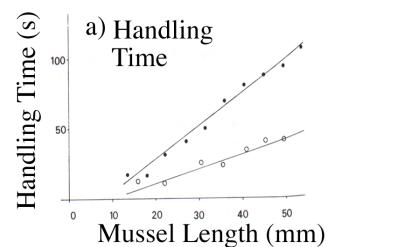
Not all systems studied to date have shown such a perfect fit to the data. Indeed, when a lack of fit is observed, it may be the case that factors not considered may influence animals in nature. It is invariably assumed that animals maximize some currency, however, the maximization of this currency is subject to various constraints such as time and energy. Identifying the optimal decision rule that maximizes the currency while the animal labors under constraints is the primary goal of optimal foraging theory. Model building for Oystercatchers is detailed in Side Box 6.1. The model building process underlying optimal foraging theory entails the identification of three parameters (Krebs and Kacelnik 1991):

- The foraging **currency** maximized by both crows and oystercatchers is energetic efficiency or net energy gain/unit of time. In the examples presented below, the currency may be quite different depending on the specific needs of animal. For example, a foraging parental starling is not just concerned with caring for its own needs, but must also tend to the needs of its developing chicks. Similarly, the foraging bee could be maximizing its own efficiency as a worker, but a more likely possibility is that the bee is maximizing efficiency for its colony.

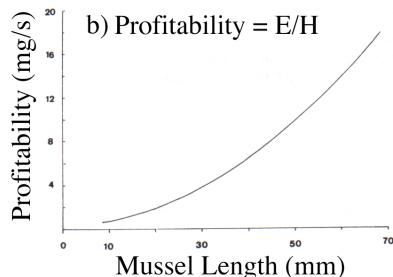
Side Box 6.1. Constraints on Optimal Foraging

The mechanics of the optimal modeling process are well illustrated by Meire and Erynnick's (1986) observations of foraging oystercatchers. The energy content of a mussel increases roughly to the cube of length (Dry Weight (mg) = 0.12 length^{2.86}) and large mussels have an enormous pay-off relative to small mussels. The following temporal, energetic, and ecological constraints set limits on decision rules adopted by oystercatchers.

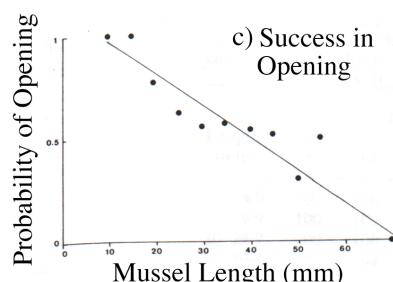
a) Is the pay-off for large mussels offset by the increased handling time? The oystercatcher's handling time increases linearly with mussel length for both the mussels that they open (solid dots) or those that they abandon unopened (open dots).



b) Model I: When we consider the increase in handling time for large mussels, profitability of Mussels still increases with Mussel Length (mm). Profitability = E/H, reflects energy gained per unit of handling time (Krebs 1978). Model I implies the largest mussels are always most profitable.

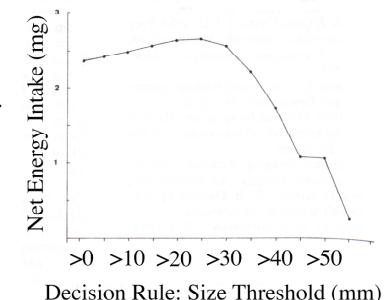


c) However, the probability that an oystercatcher successfully opens a mussel declines inversely with mussel size. An oystercatcher can open every mussel that is below 15 mm in length, but success declines rapidly as size increases and oystercatchers can't open mussels greater than 70 mm.



d) Model II: A more realistic model would adjust the profitability of a mussel by the size dependence of: energy content (E), probability of opening (P, from panel c) or failing to open the mussel (1-P), the handling time for opened mussels (H) and time wasted on unopened mussels (W):

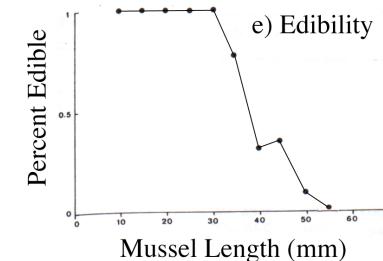
$$\text{Profit} = \frac{E \times P}{H \times P + W \times (1 - P)}.$$



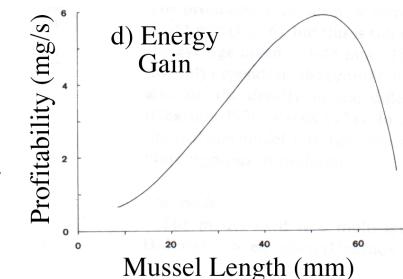
The optimal size (peak on the curve) predicted from this model is 52 mm, which is far greater than the observed 25 mm threshold.

e) In addition, mussels that are covered in barnacles are not as attractive to oystercatchers. The largest, oldest mussels have more barnacles.

However, Model II is still inadequate as it is based on the profit from single individuals, not profit that an oystercatcher can extract from foraging sequentially on the mudflat for mussels that vary in size.



f) Model III: As in seen in the example with finicky crows, if an oystercatcher rejects too many small mussels, travel time to the next suitable mussel is greatly increased. The increase in travel time causes a decrease in average profitability of being too finicky and feeding on large mussels. This shifts the curve for model II to the right. The resulting profit curve yields an optimal size-threshold for feeding of 25 mm (peak on the curve), which matches observed oystercatcher selectivity quite precisely (Fig. 6.3).



ii) Foragers also work under energetic and time **constraints**. The time **constraints** may be fixed, as in the case of crows, which have a constant time to find the next item irrespective of prey size. Alternatively, the constraints such as handling time may vary with prey size, as in the case of oystercatchers (see Side Box 6.1). The energetic costs of foraging activities such as flight and walking vary enormously. Failure to identify all constraints, and the precise nature of the constraints will result in a model that has poor predictive power. Even in the case of a simple model for foraging, a suite of factors limits oystercatchers, which must all be considered to achieve a close fit between theory and observation (see Side Box 6.1). Finding the constraints may entail an iterative process; the complexity of an optimal foraging model is gradually increased and constraints are added until all salient ones have been identified and good fit is achieved.

iii) The appropriate **decision rule** must also be identified. A test of optimal foraging compares the observed size threshold with that predicted from the size distribution of prey in the environment and the constraints of foraging. The observed threshold size for acceptance of prey items for crows and oystercatchers appeared to match the predicted threshold size quite closely indicating a good fit with the model.

Richardson and Verbeek (1986) only considered a single model of optimal foraging. Animals labor under time and energy constraints that are independent of prey size and additional ecological constraint relates to the rarity of the largest, most-profitable prey. Meire and Ervynck (1986) considered three different models of optimal foraging that varied in the number of constraints built into the model. The simplest model only included the profitability as a function of prey length. A more complex model factored in the difficulty in opening prey of various size, and the attractiveness of prey (e.g., barnacles make mussels more difficult to open). The most complex model also factored in the availability of mussels on the beach. The simplest foraging model did not adequately predict the observed size threshold of acceptance, nor did the second model, but a more complex model that included the increased handling time and difficulty of cracking thick-walled mussels provided a surprisingly good fit to the observed size threshold. It is often the case that behaviorists first consider the simplest model before proceeding to a more complex explanation for the behavior of animals.

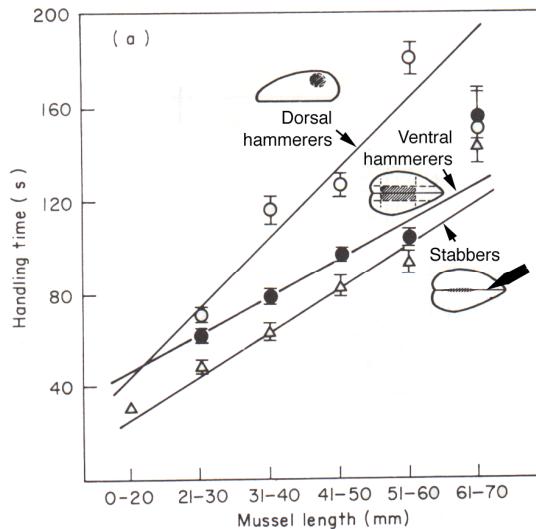
Finally, the crow and the oystercatcher faced the same basic search constraints. The size distribution of prey in the environment was a major factor governing whether or not a bird accepted or rejected a prey item. The size distribution of prey is an example of how ecology of the prey constrains the optimal foraging solution adopted by the birds. It is not necessarily the case that mussel availability remains constant throughout the year. Moreover, not all animals use the same foraging strategies, as there is more than one way to crack a nut. Crows drop clams while oystercatchers hammer them open. Individuals within a single species might likewise vary in their use of alternative feeding strategies.

Variation in feeding mechanisms within a population

Our models of crows and oystercatchers suggest that there is one unique decision rule that maximizes energy intake per unit of time. However, animals vary dramatically in the kinds of foraging behaviors that they use in nature. Differences in foraging techniques can have a dramatic effect on the optimization decision rules that various individuals use in a single population. Cayford and Goss-Custard (Cayford and Goss-Custard 1990) have observed oystercatchers foraging with three styles:

1. stabbers that use their bill to stab the vulnerable area between the valves,
2. dorsal hammerers that use their bill to hammer through the dorsal surface of mussels, and
3. ventral hammerers that opt for the opposite side.

Each foraging style has different handling times. Dorsal hammerers take the longest to break through the mussel followed by the ventral hammerers. The stabbers are the fastest at cracking mussels open with their bills. Given this efficient style, stabbers should feed on the largest mussels. Conversely, dorsal hammerers should feed on mussels that are intermediate in size. These gross expectations are borne out by natural history observations made by a number of researchers (Norton-Griffiths 1967; Ens 1982). Every factor considered by Meire and Ervynck (1986) to be constraints on foraging oystercatchers (see Side Box 6.1) were also found to differ for the individual oystercatchers that adopted one of the three feeding-styles (Cayford and Goss-Custard 1990).



→ Figure 6.4. Handling times observed by Cayford and Goss-Custard (1990) for oystercatchers feeding on mussels with three different styles: dorsal hammers, ventral hammers, and stabbers. Attack point used by oystercatchers on mussels are shown.

In addition, the availability of mussels in different size classes changes on a seasonal basis. Oystercatchers should be able to accommodate changes in the sizes of mussels available, by flexibly adjusting their decision rule. During the winter large mussels are eaten or removed from the shore by heavy wave action, and newly settled mussels of small size gradually replace them. The smallest mussels are found in the spring during March and April. Cayford and Goss-Custard (1990) used an optimal foraging model that was fashioned from the same kinds of constraints that were used by Meire and Ervynck (see Side Box 6.1).

Additional observations were made on the changes in the availability of mussels as a function of season, as well as the variation in handling times for the three feeding styles. A very good fit was found between the predicted decision rule and the observed size-selectivity of oystercatchers during most of the year. A lack of fit between observed and predicted selectivity was only found for dorsal hammerers during the winter months (5 time points), but good fit was still observed for 5 time points. The fit was excellent for ventral hammers for 9 out of the 9 time points they considered.

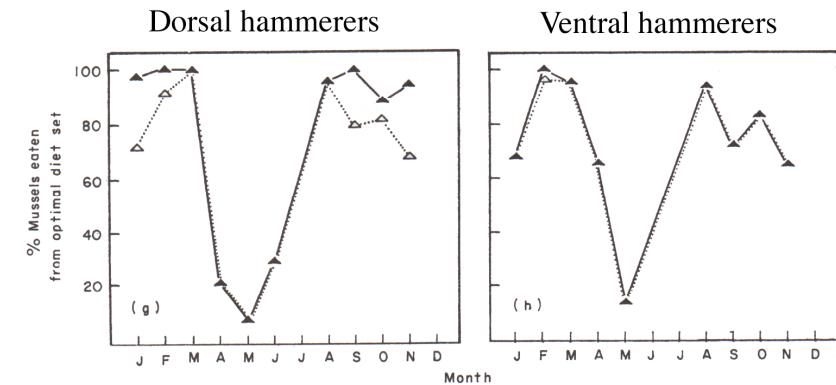


Figure 6.5. Observed and predicted size selectivity for two oystercatcher feeding styles across the year. From Cayford and Goss-Custard (1990).

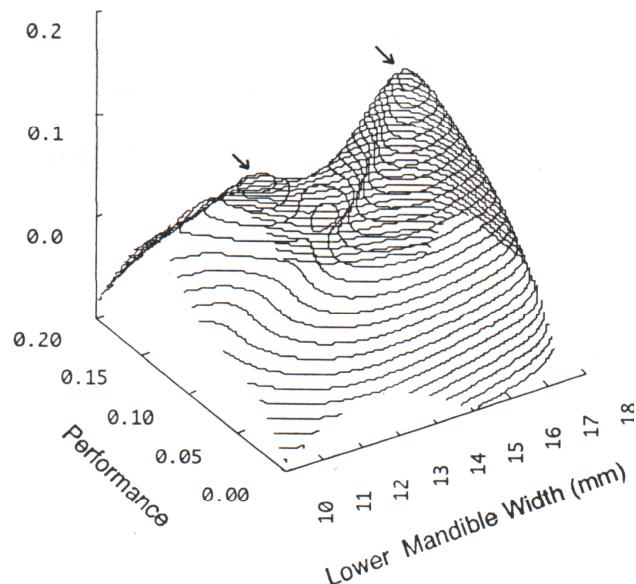
Individual foraging determined by genes, environment, or culture

The differences in foraging styles seen in oystercatchers are culturally transmitted. Parental oystercatchers teach their chicks how to forage by taking them out into the intertidal (Bruno Ens, personal communication). Chicks learn the distinctive style from their parents and the chicks in turn pass the style on to their own offspring. Many birds and mammals pass on their foraging strategies to offspring. In some cases, foraging styles are distinctive among primate groups and have a strong cultural basis (Lefebvre 1995), as do certain novel foraging tactics in birds, such as the ability to open milk bottles (Lefebvre 1995) (covered in more detail in Chapter 19: Societal Evolution).

Cultural transmission of foraging behaviors contrasts with genetic differences in foraging behavior. Different foraging behaviors are constrained by morphological differences among individuals in a population. For example, the foraging behavior African seed crackers, *Pyrenestes ostrinus*, are constrained by their genes and morphology because a seed cracker is born with a small or large bill depending on the alleles at a single genetic locus (see Chapter 2). Each bill morph has very distinctive foraging preference that has a profound effect on the profitability of feeding on different seed sizes from the various species of sedge that are used as a primary food source by the birds.

Moreover, the curve describing a bird's probability of survival closely matches the curve describing a model for optimal foraging (Smith 1987; Smith and Girman 1999) (Fig. 6.6). Natural selection is constantly refining the birds into two distinctive types by eliminating birds with intermediate-sized bills and thus a poor rate of energy gain while feeding on the seeds. Not all birds are well adapted. Whereas small seeds and large seeds are quite abundant, there are very few medium-sized seeds in the environment. If large billed birds can eat all seeds, regardless of size, why are there still small-billed birds? Large-billed birds cannot eliminate the small-billed morph from the population through competition because large-billed birds face handling time problems of small seeds. Small birds can handle small seeds quite well, but cannot handle large seeds. The converse is true for large-billed birds. Those unfortunate to have been born with an intermediate-sized bill feed poorly on all seed types.

Figure 6.6 Survival of African seed-crackers, *Pyrenestes ostrinus*, depends on bill size and feeding performance on sedge seeds (Profitability) of different size. The effects of bill size on survival mirror the effects of performance on survival. Two "fitness optima" correspond to large morphs feeding on large seeds and small morphs feeding on small seeds. Birds with intermediate bill size and performance have poor survival (from Smith 1997).



Many animals have discrete differences in morphology within a single population like the African seed crackers. Such genetically based differences have a profound effect on foraging behavior and can ultimately lead to new species. Lake-dwelling Arctic char, *Salvelinus alpinus*, have repeatedly evolved several different morphs with dramatic differences in body shape and morphology in the same lake (Skulason et al. 1989). The morphs are associated with strong preferences for prey that are typically found in different habitats of the lake. Fish that forage for plankton in the open-water tend to have long streamlined bodies, while those that favor inshore foraging on bottom-dwelling prey have shorter compact bodies (Schulter 1996). The relative foraging efficiency of each type tends to be highest when feeding on food items to which they have morphological specialization (Skulason and Smith 1995; Schulter 1996).

Discrete differences in behavior and morphology do not need to have a strict genetic cause. The environment can trigger changes in morphology and behavior during development. Individual larvae of spade-foot toads, *Scaphiopus couchii*, in the desert southwest can change into the omnivore morph or the carnivore morph depending on the availability of shrimp (see chapter 2, 8). Thus, the differences in foraging behavior are induced by a difference in the environment. Carnivore tadpoles tend to forage by themselves whereas omnivores tend to forage in large gregarious schools. Carnivores tend to have a much larger jaw muscle and keratinized tooth, while omnivores tend to have a longer gut. Each morph grows and metamorphoses to large size only if they are foraging on the food to which their behavior and morphology is specialized -- carnivores metamorphose at a large size on shrimp, but small size on detritus, whereas the converse is true for the omnivore morph. The larvae of the spade-foot toads have no a priori information regarding the presence of shrimp in a given pond. If shrimp are common in their diet then it is beneficial to be able to efficiently exploit this resource and metamorphose quite rapidly. When shrimp are rare or absent, the larvae should feed with the alternative behaviors and preferences of an omnivore.

Plasticity in foraging behavior and morphology serves the spade foot larvae well. Behavioral and morphological plasticity as a function of environmental conditions reflects a 'strategic evolutionary alternative' to

the hardwired genetic coding of behavior and morphology under a system of genetic control. Cichlid fish show similar capacity to alter their behavior and morphology. When hard prey such as shelled snails are available, feeding on snails causes a special second jaw, the pharyngeal jaw, to become much more ossified and the muscles become greatly enlarged in the Midas cichlid (Meyer, 1988). The behavioral act of masticating triggers dramatic changes in morphology and the changes in morphology have a feedback effect on behavior. In the absence of hard food, the fish maintain a much less ossified pharyngeal jaw. Bluegill sunfish can also change morphology during juvenile development in response to prey hardness (Ehlanger and Wilson 1988). Such early development events are triggered by the type of prey found in the environment and have a long lasting effect on foraging behavior and efficiency throughout adult life.

In summary, the foraging behavior can change during the course of a single season as individuals adjust their decision rules. Foraging behavior can be transmitted culturally across generations as parents teach offspring their own predilections of foraging. Genetic effects can also dictate foraging behavior. Animals might be born with morphologies, which are dictated by the genes, and the morphology constrains their foraging behavior for life. Alternatively, morphology can change during early life or adult life depending on the kind of food available for feeding, the foraging environment. Such changes can be reversible, but more often than not the changes in morphology are irreversible.

The Marginal Value Theorem and Optimal Foraging

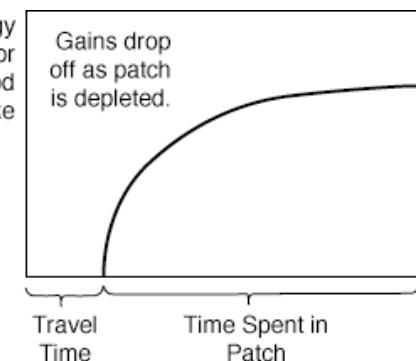
In most of the previous examples, individual prey items varied dramatically in quality, but prey was scattered randomly about the environment. Prey may also be found in relatively discrete patches. When prey items are found in a patchy distribution, the predator must expend considerable energy traveling from patch to patch. The choice about when, where, and how long to settle to feed is another one of the basic decision rules for an organism searching for resources among widely scattered patches. One of the simplest solutions in foraging ecology, the **marginal value theorem** (Charnov 1976), is easy to derive from a simple graphical analysis.

The marginal value theorem specifies the "giving up time" or when an organism should leave a patch that it is exploiting. As an animal begins to feed, its energy gain gradually begins to slow down when food becomes scarcer in the patch. It takes longer and longer to find the next item of food. The marginal value, or amount of energy remaining in the patch declines as the patch is exploited. The curve describing **energy gain** as a function of the amount of time a predator spends in a patch, starts off with a steep slope that gradually levels off as the prey becomes depleted. Eventually, if the predator stays in the patch long enough, all food is consumed and no more energy can be gained.

The marginal value theorem has broad applicability to many optimal solutions in animal behavior (Krebs 1978; Krebs and Davies 1987; Krebs and Kacelnik 1991), and as we shall see in upcoming chapters the theorem has been applied to behaviors as diverse as territory defense, mate search, and sperm transfer. In the case of a feeding bird that has found a patch, items may be consumed immediately. The depletion of the patch results directly from the removal of prey. In the case of a parent foraging for its young, items are collected and transported as a load back to the nest to feed the chicks. As a parental bird loads up on items, efficiency at collecting the next item declines with each item that is stuffed into its bill. In both cases, the marginal value of the patch or marginal value of loading up on an additional prey item declines with increased time spent in the patch.

Marginal Value Theorem: Travel Time and Energy Gain

When should the animal give up on a patch and move on to find a new one? The crucial parameter that governs this decision is not only the amount of time spent in the patch, but also the travel time between patches. An animal gains no energy while traveling, and in fact expends considerable energy during locomotion between patches. Thus, the value maximized by the forager



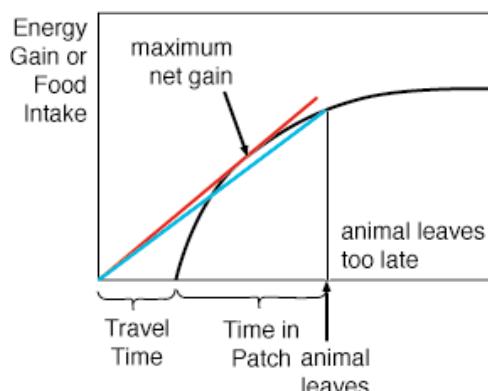
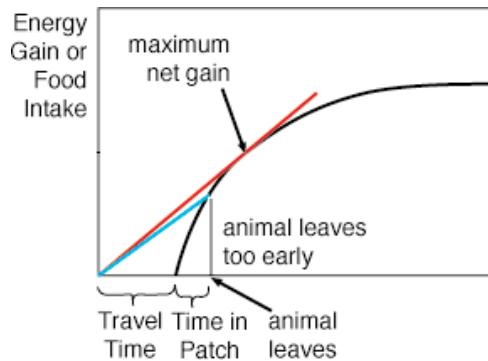
should be net rate of energy gain, which includes time during which it cannot feed as it travels to a patch:

$$\text{Rate of Energy Gain} = \frac{\text{Energy}}{\text{Time}} = \frac{\text{Energy Gain or Load Size}}{(\text{Travel time to patch} + \text{Foraging Time in Patch})} \quad 6.1$$

The rate of energy gain is expressed as calories gained per unit of time, which on the graph at the right corresponds to the slope of a straight line, the rise over run or Energy Gain/Time. A steep line (gain/time) or the line of greatest slope that still touches the curve will maximize the rate of energy gain. An animal that leaves too early gains less energy (shallow line) relative to the maximum net gain that is possible.

There is really no benefit in staying too long once the tasty treats begin to run out. The animal should move on to greener pastures. Consequently, an animal that leaves too late also has a shallow line relative to the line of maximum slope. The line that gives the maximum rate of energy gain is the line that hits the gain curve at a **tangent (red line)**. The tangent is a line of steepest slope, which intersects the gain curve at a single point.

Finally, animals should also be sensitive to the length of time it takes to travel from patch to patch. When the travel time is short, they should leave far sooner than if the travel time between patches is very long.



Currency and Load Size in Parental Starlings

European Starlings, *Sturnus vulgaris*, have provided Alejandro Kacelnik with a model system to test key assumptions underlying optimal foraging -- the units of currency used by animals when they forage in discrete patches (Kacelnik 1984). Starlings fly from their nest to feeding sites, which may be located at a variety of distances from the nest. The travel time for the parental starling is given to be the **round trip time** from the nest to a food source and back to the nest to feed the young. When the starling arrives at a foraging site such as a patch of grass in the forest, it begins probing the soil and spreading its bill to expose larval leatherjackets (*Tipula* spp.), which are considered to be a delicacy among starlings. Each larva is extracted from the soil and placed near the juncture between the mandibles of the bill.

A starling is capable of holding several larvae with its mandible, while simultaneously extracting other larvae from the soil. However, the speed with which the starling extracts larvae declines with the number of larvae that it has already stuffed into its bill. The parental starling should experience the force of the marginal value theorem. It should return to feed its young before it begins to bobble and drop larvae. The marginal gains from remaining on the patch and feeding will be offset by the increased time wasted fumbling with larvae.

How many items should a starling parent collect in its bill before returning to feed its young at the nest? Rather than study starlings foraging for natural prey in the patches of grass between forests and hedgerows around Oxford University, Kacelnik (1984) was able to train some birds to visit a feeder -- an artificial patch. At the feeder, Kacelnik simulated the marginal gain curve a foraging parent might experience while feeding on natural prey. The feeder dispensed mealworms at an ever-decreasing rate once the parent bird arrived at the feeder. The ability to adjust the gain curve allowed Kacelnik to give his test subjects a constant gain curve throughout the duration of the experiment. Under natural conditions, this would be far harder to accomplish because random factors, such as the distribution of prey in the soil, might add uncontrolled noise to the expected gain curve. In addition, Kacelnik could manipulate the round trip travel time by moving the feeder further and further from the nest. The parental starlings adjusted readily to the movement of the feeder.

Kacelnik modeled the time constraints and the decision rules regarding the number of mealworms the parent collected before returning to the nest. He was particularly interested in exploring alternative currencies, which could be used by a parental starling. Up to this point we have only considered currency expressed as profitability per unit of time. Many other currencies are possible. For example, starlings might be using a currency based on efficiency:

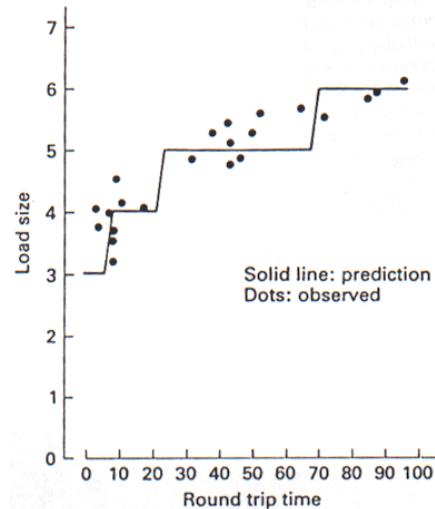
$$\text{Energy efficiency} = \frac{\text{Energy Gained}}{\text{Energy Spent}}, \quad 6.3$$

which maximizes the energy return for energy invested. Alternatively, parental starlings might be maximizing the energy delivery to progeny per unit time:

$$\text{Energy delivery} = \frac{\text{Energy per load} - \text{Parental energy costs} + \text{chick energy costs}}{\text{Round Trip Travel Time}}, \quad 6.4$$

which entails the energy costs of the parent during travel, as well as the energy costs for the chicks between deliveries. Energy delivery as a currency would maximize the delivery rate of prey to the chicks. Notice that the energetic efficiency does not consider time (6.3), but time is factored into the denominator of the delivery equation (6.4).

Figure 6.7. Observed load sizes for parental starlings, *Sturnus vulgaris*, foraging on mealworms at feeding stations located at different distances from the parents' nest. The stepped line describes the load sizes that are predicted from an optimal foraging model that is based on a parent that maximizes the amount of energy delivered to chicks back at the nest.



A few basic predictions can be made to test whether the marginal value theorem applies to starlings (see Side Box 6.3). When the feeder is very close to the nest, the parental starling should wait for fewer prey items before it flies back to the nest to feed its chicks, compared to a feeding station that is located farther away from the nest.

The idealized pattern for the gain curve would be a step-function (Fig. 6.7) in which the height of the step is a constant expressed as a unit prey item, the mealworm. A parent can only deliver mealworms in increments of one. The length of each step is the release rate of mealworms, which was adjusted experimentally by Kacelnik at the feeding platform. Parental starlings settle on an observed energy gain function that is a close match to the line predicted with the model for energy delivery rate to chicks. Lack of fit between observed and predicted data would be reflected in the points above or below the predicted line. Points found above the predicted line reflect situations when parents take more prey than expected by the optimal foraging model. Points below the predicted line represent the case when parents leave with far lighter loads. The scatter of points is tightly clustered around the theoretical curve for optimal foraging based on maximizing energy delivery to young.

The currency used by adult birds is only slightly different from that used by crows and oystercatchers. Based on these observations we would predict that animals maximize gain per unit time when feeding themselves (e.g., data from crows and oystercatchers) and maximize gain for the chicks when nesting (e.g., data on starlings). The needs of the young supersede the needs of the parent. One might expect that the starling might adjust the number of items that it carries per flight if some other factors constrain the number of flights. If the parent faced additional costs during flight that are less tangibly expressed in long-term physiological cost, the starling might alter its currency. If the length of flight were directly related to a parent's lifespan, then the currency they use might reflect the long-term consequences of flight distance. Unfortunately, the long lifespan of starlings makes this proposition difficult to test. Animals with shorter lifespan, such as insects, provide us with more suitable model systems to test hypotheses of the long-term physiological constraints on foraging and the effect of such constraints on the currency.

Currency of Bee Colonies

Bees forage for honey and pollen at flowers located at a variable distance from the nest (Heinrich 1978). The load of honey or pollen is returned to the nest where other workers store it in honeycombs, to be used later in feeding developing young. Like starlings they are central place foragers returning to a central location to deposit their load. It seems intuitively appealing that highly successful colonies bring in larger amounts of energy per unit time as a function of the number of workers. Colony growth and output of progeny should be directly related to the energy gain of the colony. Because the food resource of bees is made up largely of nectar (with some pollen), it is very easy to quantify colony success in terms of the rate of energy gain per unit time.

What factors determine the rate at which colonies gain energy? Certainly the availability of flowers and distance of flowers from the colony affects foraging effort. Flight in bees is a costly enterprise (Heinrich 1978). How far from the nest should a worker fly before pay-offs from distant flowers become unprofitable? Taking on a load of nectar or a load of pollen makes flight for the bee even more costly. Given the load must be returned back to the colony, flying with a load costs calories compared to unloaded flight. How much pollen and nectar should a bee take on board before heading back to the hive? All of this work leads to wear and tear on busy worker bees. Work is hazardous to health and there is no workers compensation pay for bees, no retirement; debilitated bees die.

What counts in the economy of a bee colony: maximizing the rate of return per unit of time, or maximizing the rate of return per worker (Schmid-Hempel et al. 1985; Schmid-Hempel 1987)? The dichotomy between these two currencies emphasizes the different constraints operating on animals. If there were no *long-term* consequences of a particular foraging strategy, then increasing the *short-term* return would benefit the colony. If, however, workers are a commodity and rearing a worker is costly (and when a worker starts working, foraging is quite costly) then lengthening the foraging lifespan of a worker might be the optimal strategy for a colony. If so, an individual worker might be expected to maximize the amount of energy collected relative to the energy used in physiologically expensive activities such as flight. Schmid-Hempel's test for honeybees (1985), like Kacelnik's (1984) test

of starlings, computed two different currencies that might be optimized in bees: 1) maximize energy intake per unit time (or a rate maximization optimal foraging model), and 2) maximize efficiency or energy collected/energy expended.

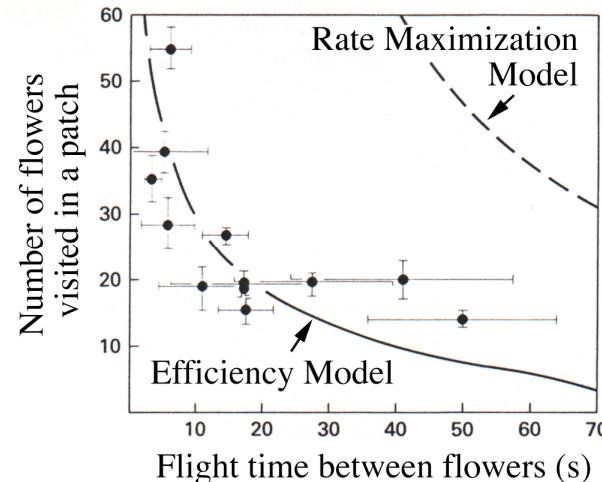
To test each of these models Schmid-Hempel collected bees that were trained to forage in artificial flower patches with a constant 0.6 μl of nectar at each flower. Bee flight is quite costly and the costs rise as a function of distance. Therefore, Schmid-Hempel varied the flight cost of bees by changing the spacing between the main flower patches. The travel time to the main flower patch was a constant distance from the hive. When flowers on the main flower patch were widely spaced on the grid, the bees would have to fly further per unit reward, and thus incur greater costs of flight compared to grids with more closely spaced flowers. Under conditions of widely spaced flowers, bees should experience a diminishing return based on the number of flowers visited.

Under a rate maximization currency, if a bee visits too many flowers in a patch before returning to unload at the hive their flight time is increased while it is carrying a nearly full load of nectar. The added costs of flight with a heavy load cuts into profit because too much energy is wasted in flight. A bee using extra energy in flight is analogous to the starling parent fumbling with one more prey item despite the fact that its bill is nearly full with larvae. Airline companies are well aware of the problem of payload and efficiency and they load up a jumbo jet with just enough fuel to make the flight from Los Angeles to New York. A little more is put in if the flight is to London. Fuel is adjusted for the number of passengers. Topping a jumbo jet off with a full load of fuel would be wasteful as the fuel is expensive. Of course airline companies also add extra fuel in case the planes run into difficulty and are rerouted to an alternate destination. Bees do not face the rerouting problem but do face a similar economic problem of fuel weight and payload weight. Under an efficiency maximization model, the worker might be expected to return back to the hive with a much smaller load overall, because an excessively large load would drastically increase the flight costs on the return leg of the journey back to the hive.

Schmid-Hempel and colleagues computed optimal foraging models for the bees based on the two currencies: efficiency maximization versus

rate maximization. The observed data was not even close to the curve predicted by the rate maximization model (Figure 6.8). Rate maximization would predict that the bees should visit far more flowers on any given foraging bout to the main flower patch. Either the bees are lazy or they are maximizing a currency based on efficiency. The match between the observed data and a model based on efficiency maximization was very good.

Figure 6.8. The number of flowers a bee visits in a patch is expected to drop off as the flight time between flowers in a patch increases, largely because of the cost of flight increases as the bee becomes loaded. The theoretical curve for a model of efficiency maximization nearly matches the observed bee behavior, while a model based on rate maximization provides a poor fit. (Schmid-Hempel et al. 1985).



In a second experiment, Schmid-Hempel and Wolf (1988) used weights to provide an ingenious test of the survival costs predicted under the efficiency model. A less efficient bee is expected to incur a reduced lifespan. By adding weights to the bees he reduced lifespan from 10.8 days in controls to 7.5 days in bees with the extra 20 mg that was glued to their abdomen. Bees pay for foraging in terms of a physiological cost of lifespan.

Ralph Cartar (1992) provided another ingenious field test of the survival consequences predicted under the foraging efficiency optimization hypothesis. Rather than add weights to the bees, Carter clipped their wings at the margins. By clipping their wings, he increased the foraging costs in a manner analogous to the costs incurred with added weight. He

made them less efficient fliers on their foraging trips away from the hive. Clipping their wing margins led to a higher rate of mortality (Fig. 6.9). In addition to the experimental survival costs of foraging that Cartar was able to induce by wing clipping, it is equally important to determine if such costs act on unmanipulated bees. Cartar correlated natural longevity with degree of wing wear. Wing wear naturally accumulates during a bee's lifespan and thus a young bee, which has longer to live, will have less wing wear than an old bee. Cartar found the expected negative correlation between natural wing wear and survival -- bees with less wing wear had more days of life left than bees with more wing wear.

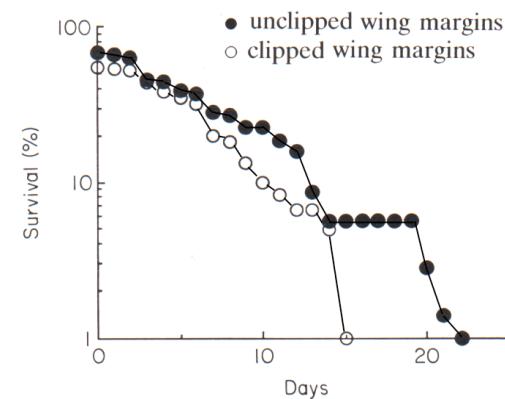


Figure 6.9. Survivorship of bumblebees, *Bombyx* spp., foraging with intact wing margins is higher than those bees laboring under the additional flight costs that are induced by clipping wing margins (from Carter, 1992).

Both experiments on survival costs support the notion of physiologically based foraging costs in bees. Long-term fitness costs can dramatically alter the apparent currency that animals use in the short-term. In the case of starlings, the loads that they carry do not appear to incur survival costs. The burden of worms they carry back to the nest is a trivial cost, relative to the overall cost of flight. For bumblebees, which have a shorter lifespan, and carry a larger load of nectar relative to their body weight, the physiological costs and wear-and-tear of loaded flight are large enough to alter the currency used during their bouts of foraging. Other studies have shown that Lapland longspurs appear to minimize flight costs as they shuttle food back to their chicks (McLaughlin and Montogomerie 1985; McLaughlin and Montogomerie 1990). Why should a long-lived bird be so affected by costs of short flights to and from the nest?

A general principle is emerging for species in which the transport costs

of foraging are substantial and when foragers are provisioning progeny (e.g., Lapland longspurs parents) or the colony (e.g., bees). When the daily delivery of energy is constrained by relatively expensive flight that requires costly self-provisioning (e.g., feeding during the day), maximizing efficiency of flight will likely be the strategy that maximizes food provisioned to either the chicks or to the colony (Ydenberg et al. 1992). This is because the cost of transport is an enormous proportion of the daily energy budget and any savings in transport can be used to increase the total amount delivered to the nest or colony. In contrast, under situations where self-provisioning costs are modest (e.g., parental starling), rate maximization appears to be the optimal strategy.

The difference between the currency used by various animals also relates to the issues of the levels of selection discussed in Chapter 4. In the case of foraging crows or oystercatchers discussed above, selection acts on the level of the individual. Maximizing energy per unit time maximizes individual fitness. In the case of foraging parental starlings, maximizing energy delivery to nests again maximizes individual fitness, because the parent maximizes successful production of young. However, the foraging honeybee is a case of kin selection in which the optimum solution for the colony is best met by getting the most out of each worker. The costs of raising a worker are so high that maximizing the returns per worker will maximizes energy gains for the colony as a whole.

Foraging in the face of a risky reward

There is an aspect of human behavior that seems quite puzzling (at least to many of us) and that is the addiction to gambling. Why should a gambler pour immense quantities of money into a business like a casino that is designed to make a profit? To be sure, the casino won't be loosing in the long run. Money isn't magically created at casinos, so the casino makes a profit at the expense of the "Joe-average" long-term loser. Many humans are **averse to risk** in that they avoid most risky situations and relatively few of us gamble on a regular basis. We tend to shy away from the "quick-pick lottery ticket" to the high-life and opt for a more steady and low-key source of income.

Most animals appear to be similarly conservative in that they tend to

avoid risky situations. The most immediate kinds of risk that animals face is the risk of starvation. The idea that animals are averse to risky rewards arose from some pioneering experiments by Caraco and his colleagues on yellow-eyed juncos, *Junco phaeonotus*, which is a common songbird of the forests of North America (Caraco et al. 1980; Caraco 1981; Caraco et al. 1990). The experiment is deceptively simple. They trained birds to feed at two kinds of feeders. One of the feeders dispensed a constant reward of 3 seeds with every visit. The other feeder dispensed a reward of 0 seeds for half of the visits and 6 seeds on the other visits. The variable reward feeder would randomly disperse 0 or 6 seeds on any given visit. Thus, it might be possible to get a string of 0's or a string of 6's as a reward. Given a choice between these two types of feeders, which give on average an identical reward of 3 seeds, the juncos opted for the less risky feeders.

The behavior of juncos is termed **risk-aversive foraging** in that they opt to feed at a station that supplies a constant rate of food and avoid feeding at a station with a variable amount of food. The level of risk-aversion of juncos can be "titrated" by gradually increasing the mean value of the variable feeding stations relative to the value of the constant-reward stations. By increasing the variable reward relative to the constant reward you can determine how much more valuable the risky feeding station must be to attract juncos. At some point, the preference of juncos for constant-reward stations will disappear and they will opt for the constant reward station versus the variable-reward stations in a 50:50 ratio. It takes nearly twice as much food at risky stations before birds begin feeding at those stations with a frequency that equals their use of the constant-reward station. At face value, such pickiness in favor of constant rewards would appear to fly in the face of optimal foraging reason. The paradox of their behavior is that the birds could do far better by opting for risky feeding stations that have twice the average pay-off, and yet they still opt for the constant reward stations approximately half of the time.

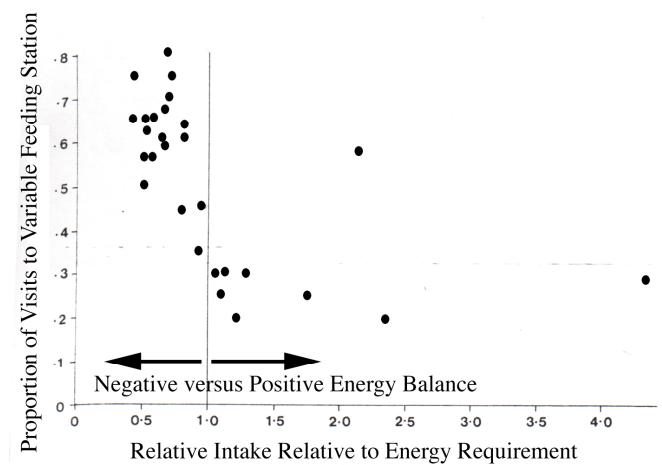
Food limitation and risk-averse vs. risk-prone behavior in shrews

To resolve this dilemma, let us return to the foraging problems faced by the common shrew, *Sorex araneus*, which were posed at the outset of this chapter. Recall that shrews have an unusually high energy demand (Vogel 1976) owing to their small size (a few grams), high surface area-

to-volume ratio and thus greater heat loss compared to a larger animal with comparable body shape and insulation capabilities (e.g., a rat) (Peters 1983; Calder 1984). Shrews face minute-by-minute decisions that affect its survival (Barnard and Hurst 1987). Because the shrew is always in a state of hunger, we might expect them to be quite choosy about their food source. Perhaps the common shrew is risk-prone in that it is more willing to opt for risky sources with a higher payoff.

Barnard and Brown (1985) introduced shrews into small tanks, to which they readily acclimated and built a nest on one side of the tank. Two feeding pots were set up on the other side of the tank. At the outset of the experiment, Barnard and Brown (1985) measured the ad libitum feeding rates, or voluntary quantity of mealworms consumed by each shrew when mealworms were available in excess quantities in both feeding pots. For each shrew, they placed mealworm chunks in each pot at constant rates (1 chunk /visit) versus variable rates (2 chunks/visit half of the time and none the other half of the time). They also tested each shrew under different overall rates of feeding by removing the stations. When the stations were periodically removed (e.g., both stations removed half of the time), the rate of intake was below the daily requirement of shrews compared to the previously described regime, which was above the daily energy requirement of individual shrews.

Figure 6.10. Proportion of visits of foraging common shrews, *Sorex araneus*, to two feeding pots that differed in the variability of reward. The frequency that shrews used variable-reward pots vs. constant-reward pots was influenced by hunger state of From Barnard and Brown (1985).



When shrews were on starvation rations, and the total amount of energy delivered fell below the minimum energy requirement, the shrews avoided the constant-reward feeding pot in favor of the variable-reward pot (Fig. 6.10). When the shrews were on ad libitum rations they opted for the constant food pot. Shrews are **risk-prone** and feed on a variable food source, but only when they are low on energy reserves. In contrast, they become averse to risk when they have abundant food.

Adaptive value of risk-sensitivity and the threat of starvation

This striking switch in behavior from risk aversion to risk prone foraging behaviors when animals face the threat of starvation is common. When an animal modifies its choice between a risky food versus constant food patch depending on their physiological state, the animal is considered to be **risk sensitive** in its foraging behavior. As seen earlier, Caraco demonstrated that well-fed juncos are similarly risk averse. In a suite of experiments on juncos, Caraco and his colleagues supplied more experiments documenting risk sensitivity. If you deprive juncos of food causing them to be energy limited, they switch from risk averse foraging, to a mode that tends to select variable, but more profitable rewards. When juncos are energy-limited to the point of starvation, they are pushed to adopt the risk seeking strategy that may yield higher payoff. Environmental cues that alter risk-prone and risk-aversive behavior in animals need not be starvation and an empty belly. Any environmental factor that correctly predicts future energy demands could be used as a cue to alter the pattern of risk sensitive foraging. Juncos appear to be sensitive to the temperature that they experience during the day (Caraco et al. 1990). A cold day will likely mean that the night will be chilly and thus more demanding energetically. Under such conditions juncos adopt a foraging strategy that favors riskier rewards.

The logic underlying the strategy observed in the gambling shrews and juncos relates to a change in the perceived value of the reward. When food is in abundance, the value of a higher paying, but variable reward is diminished relative to a constant reward that appears to satisfy all the animal's energetic needs. However, when facing food limitation and potential risk of starvation the perceived value of a variable food source increases dramatically. It is only at risky sites that the animal can potentially find enough food to make it through the period of starvation.

Facing the possibility of starvation, animals are willing to gamble on the “strike-it-rich” policy of risk-prone foraging. If the animal were to stay on the constant reward ration, it faces certain starvation. The animal would be visiting sites that could not sustain cumulative energy needs for an entire day. It could not feed fast enough to survive. Even at a variable reward site with the same average reward, most animals would also face certain death. Some foragers will have a string of bad luck and starve. Some will have a string of average luck and still starve. However, there will always be those lucky few that experience a string of good luck. It is those lucky few that survive and pass on genes to the next generation. It seems then that under some conditions adopting a risk-prone gambling strategy could have survival value. Under food plenty conditions, risk is not rewarded as both strategies provide sufficient food.

Are animals risk-sensitive in nature?

The pattern of risk aversion appears to work for animals foraging in the context of a laboratory experiment. Such experiments are powerful because they suggest that animals might have the cognitive machinery to make relatively complex life and death changes in behavior as a function of their nutritional state. Tests of such theory in the wild are far more challenging because the amount of reward is difficult to control in nature. The distance between patches is also a difficult factor to control. It is always a pleasant surprise when someone develops the first experimental test of an interesting problem in the wild that was first elucidated in several carefully controlled laboratory experiments. In 1991, Ralph Cartar supplied a field test of risk-aversion and risk-prone switching by manipulating the food reserves in bumblebee nests.

To test these ideas in the wild, Cartar (1991) first had to find a system where two species of plant had striking differences in the variability of the nectar reward in their flowers. However, the two species had to be close enough in average reward that the difference in average payoff does not overwhelm the tendency to be risk averse. For example, the risky flowers could be so good relative to the constant reward flowers that they would always feed at the risky flower, given their overwhelmingly good return. Two common flowering plants on the west coast of North America satisfy these stringent conditions for a field experiment (Fig. 6.11):

1. seablush, *Plectritis congesta*, has low variance among flowers and a mean reward of 2015 Joules/load, and
2. dwarf huckleberry, *Vaccinium capespitosum*, has flowers with 18 times the variability in reward, but only slightly larger average reward at 2183 Joules/load.

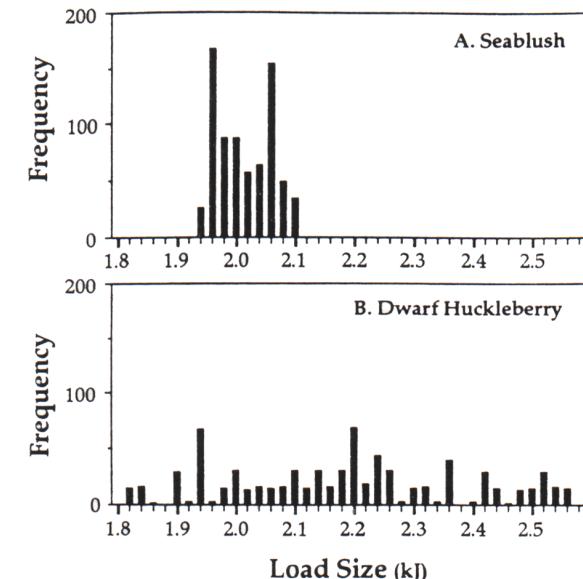


Figure 6.11. Variability in nectar load size for bumblebees, *Bombyx* spp., feeding on either seablush, which has a more constant reward than the highly variable reward of dwarf huckleberry. Bumblebees switch between seablush and dwarf huckleberry depending on the energy reserves in the colony (see text).

When Carter removed honey from the bumblebee colonies he switched the foraging preferences of the colony from the constant reward of seablush to the variable reward of dwarf huckleberry. Adding honey to larder of colonies had the opposite effect of switching the colony preferences from the risky dwarf huckleberry to constant reward of seablush. Cartar's (1991) food deprivation and food supplementation at the level of the colony generated the predicted switch between risk aversion and risk prone behavior. Bees in nature are sensitive to reward risk. In this case, it is the colony reserves, not the individual foragers reserves, which causes a switch in the behavior of the individual.

Bumblebees must have an interesting mechanism for determining how much food is in the larder, given that the stores are used to supply colony reproduction. The mechanism bumblebees use to assess the

colony energy reserves is currently unknown. However, work on honeybees (discussed in greater detail in Chapter 17: Learning) suggests that a returning forager must also ‘forage’ for a bee to unload its crop. The time taken by a returning forager to find a bee that can store its food and thus unload gives the forager data on energy flow in the colony (Seeley and Tovey 1994). A longer wait would suggest that workers that store honey are at the limits of their processing capabilities. This implies that the reserves of the colony are being replenished at a rapid rate by colony-wide foraging success.

Decision Rules and Cognition During Foraging

Cognition and Perceptual Constraints on Foraging?

By considering the proximate mechanisms underlying decision rules, it will be clear that cognitive processes may place limits on the kinds of optimal choices that animals can make. Cognitive processes in animals are defined in terms of three processes (Roitblat 1987; Real 1991):

1. **perception** -- a unit of information from the environment is collected and stored in memory,
2. **data manipulation** -- several units of information, which are stored in memory, are analyzed according to computational rules built into the nervous system (we will treat the details of nervous and sensory system in later Chapters),
3. **forming a representation of the environment** -- a complete "picture" is formed from processing all the information, and the organism bases its decision on the picture or representation of the environment.

Optimal foraging theory assumes that animals have perfect information regarding their environment. They do not. They have to learn about their environment, and they must process and remember information. While animals may not be **omniscient** or ‘all-knowing’, they undoubtedly use some form of intelligence and problem-solving skills during their foraging efforts. The cognitive sciences have a rich tradition in the field of animal psychology (Chapter 1, Chapter 17). Animal psychologists have looked at questions regarding the evolution of intelligence,

problem-solving skills, communication, and language. A basic understanding of cognition is necessary to appreciate the biology upon which foraging *decision rules* are based. Spatial memory, cognition and learning, are so central to understanding behavior that we will consider the proximate and ultimate aspects of these issues in greater detail in future chapters. Here, I highlight a few examples to illustrate how a consideration of cognition relates to potential decision rules used during foraging.

Even though many optimality models assume omniscient knowledge, such a simplifying assumption does not invalidate the predictions of optimality theory. If there is a mismatch between models of optimal foraging and the observed pattern of behavior, the mismatch might require that cognitive processing be built into the model. A consideration of cognition serves to refine the existing theory and complements our understanding of the cognitive architecture that underlies many of the decision rules discussed above. How are decision rules implemented at the level of the brain and at the level of the neural circuits that underlie behavior? To explore these ideas, let’s revisit a problem that we considered previously, risk aversion in bumblebees.

Risk Aversion and the Cognitive Bumble Bee

When bumblebees are deprived of food they switch from a foraging pattern of risk-aversive to risk-prone foraging behavior. There are a number of plausible ultimate hypotheses already discussed which seem to explain risk-sensitive foraging in bees. Gambling can have adaptive value when animals are close to the wire in terms of energy needs. Les Real has explored the alternative possibility that risk aversion may be imposed by proximate cognitive constraints of memory on foraging. Real tested for the standard form of risk-aversion in bumblebees by laying out a rectangular grid of yellow and blue flowers. Bees have keen color vision and can discriminate these colors, even if the flowers are artificial and made from cardboard. The following experimental conditions were set up in the laboratory:

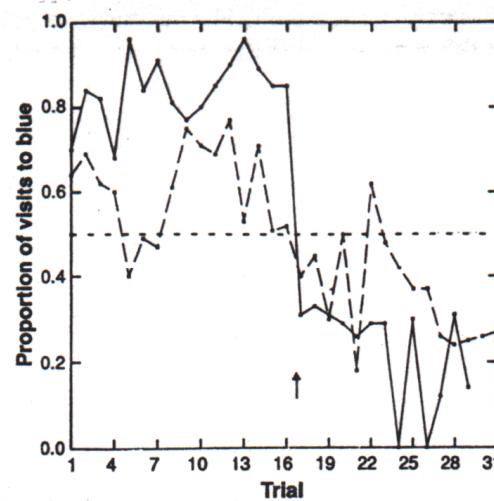
1. Each blue flower initially contained 2 µl of nectar,
2. 1/3 of the yellow flowers contained 6 µl and the remaining 2/3 of the yellow flowers contained no nectar. The average reward

across all yellow flowers would be the same as blue flowers, $6/3 = 2 \mu\text{l}$,

3. The color types of flowers with constant- or variable-reward were randomly placed in a large array -- a ‘patch’ of artificial flowers,
4. Bees were allowed to forage for half of the run with a given match between flower color and reward structure, and then the variable reward was switched for color types half way through the experiment. Blue flowers acquired the stigma of a risky payoff and yellow flowers acquired the pattern of a constant reward.

Figure 6.12. Floral preference of bumble bees, *Bombyx pennsylvanicus*, for artificial blue flowers. Solid line: The artificial patch initially contained equal numbers of yellow and blue flowers that had the same average reward ($2 \mu\text{l}/\text{flower}$), but blue flowers had a constant reward of nectar, while yellow flowers had a variable reward in each flower (0 versus $6 \mu\text{l}$). Bees are risk averse and avoid the variable yellow flowers. When blue flowers are made variable (arrow at trial 17) bees maintain risk averse foraging and avoid the now more variable blue flowers. The dashed line reflects trials in which Real used less extreme variability in reward ($5 \mu\text{l}$ in $1/3$ of the flowers, and $0.5 \mu\text{l}$ in $2/3$'s of the flowers), but, bees continue to be risk averse -- though less than the large variability treatment of 0 vs. $6 \mu\text{l}$. (From Real 1991).

The color of the constant versus variable rewards did not matter. If the constant flower was blue or yellow then approximately 70% of the visits were to constant flowers. The bees had no preference for flower color, but had a strong preference to flowers that provided a constant reward. Moreover, if they switched the reward structure during one run of the



experiment (e.g., yellow variable and blue constant to blue variable and yellow constant, Fig. 6.12), the bees likewise switched their preferences to track the less variable reward (e.g., from preferring blue to preferring yellow flowers) within a single trial.

A simple model for memory might easily explain risk aversion in bees. Imagine that a bumblebee has a specified list of items in its memory that describe profitability of items during past bouts of foraging. The number of items that an individual can recall has been termed a **memory window**. Assume that the bumblebee stores information about the last flower that it has visited in its memory such as the amount of energy it collected per unit time (Energy/Time). A sequence of such items is a memory list:

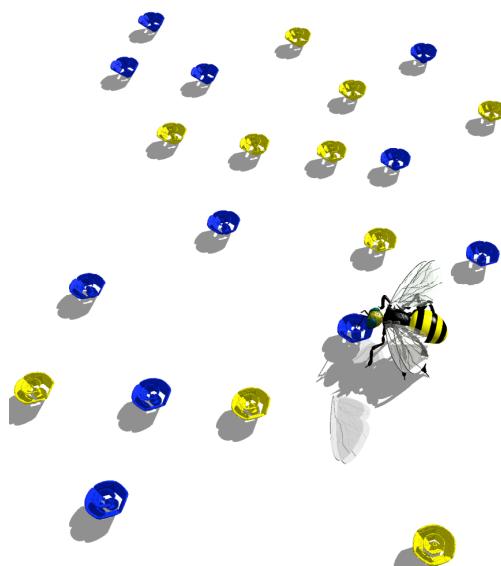
- E0/T0 (the current item)
- E1/T1 (last item encountered)
- E2/T2 (second last item encountered)
- E3/T3 (the third last item encountered)

Acquiring this information constitutes the first event in cognitive processing -- perception. If a bumblebee can only remember the last item upon which it foraged (e.g., E/T1 which was yellow) and compares this item stored in memory with the current flower (E0/T0 which was blue), then a very short memory might easily explain their aversion to risky flowers behavior. Moreover, if a bumblebee can only remember one item for each flower and then compare the last two items, there would be no need for data manipulation per se. The final phase, in cognitive processing, forming a representation of the external environment would be quite straightforward, and the bee would just compare the reward for yellow (one flower ago) versus blue flowers (the current flower). During the first few flights a bumblebee would automatically opt for constant reward flowers. Odds are 2:1 that a bumble bee will encounter 0 μl of nectar when feeding on a variable-reward yellow flower and it will always get something from blue. Two bees out of three would opt for the blue flower after this kind of sampling. By the law of probability, a third bee, might initially perceive yellow flowers as yielding a whopping reward of 6 μl . This bee would

keep sampling yellow, until it too eventually hit a 0 μl reward flower. At this point it might resample blue and compare the last two items and then invariably it would switch to blue.

Such a ‘cognitively-challenged’ bee might be risk averse if it could only remember the last item. A bumblebee feeding on the non-varying flower would perceive that it was feeding on the flower with the highest reward. It would then tend continue to seek out flowers with the same color (Fig. 6.12), but this perception of flower rewards is greatly distorted by a short memory window. If we endowed the bee with a longer list length, say two items, then things get more interesting. A few quick calculations (see Side Box 6.2: Memory List Lengths in Bumblebees), will demonstrate that a bee must have a memory window of at least 2 items for it to be capable of registering that variable- and constant-reward flowers really do have the same average reward. Even with simple models of cognition, it is clear that animals must sample their environment to learn about resources. For example, it takes a bee *at least* an average of 4 flower visits to form a somewhat realistic picture of its environment by comparing rewards from two kinds of flowers based on sampling at least two flowers of each type (see Side Box 6.3).

How many items do bees actually remember during foraging? Based on the experiments in which Real switched the reward structure during the course of the experiment (see Figure 6.10), bees may only use a single item during their foraging and sampling. Within a single trial, most bees picked up on the switch in variable flowers from blue to yellow. Bees might use a very simple decision rule: “switch flowers if nothing is found in the last flower”. However, this would imply that bees could become perfect at being risk averse and always avoid risky flowers after only a single bout of sampling. It is clear that they continue to sample some variable flowers throughout the experiment as the average preference is only 71%. Periodic sampling of alternatives may be built into the cognitive machinery of animals and variety is the spice of life.



Les Real (1991) argues that such list lengths might be adaptive and not just a constraint. You might consider it to be an ‘optimal rule of forgetfulness’. If food items are *distributed randomly* in the environment, then list lengths are of absolutely no use to a forager. Feeding at one flower provides no information about the next flower encountered. However, if the various colors of flowers are *clumped in the environment* and rewards are similar between clumped flowers, then feeding at a single flower immediately provides information about nearby flowers. The foraging bumblebee will remember the reward and use that information to decide whether to continue feeding in the clump of flowers, or leave the clump to search for higher payoffs. As food becomes more predictable, it becomes less important for animals to learn where food is abundant. When food is *perfectly predictable* in the environment, a hard-wired decision rule could be used rather than one shaped by cognitive experiences. In summary, when food is found with an intermediate amount of variability it pays to remember past experiences, and to use some sort of cognitive processing in the decision making.

How much that is kept in memory or how much information from memory is used to compute decisions really depends on the kind of variability or patchiness that an animal faces. For bumblebees, a list length of one to two items appears to be just about right. However, given the risk-prone foraging that bees adopt when food is removed from the hive (see results by Cartar), it is plausible that the size of memory list might change with the nutritional condition of animals. Hungry bees might remember more, and this could trigger risk prone behavior. Think about how forgetful you become after gorging on pizza and you have a first hand appreciation of memory and satiation.

Side Box 6.2. Memory List Lengths in Bumblebees

How large a memory window does a bee need to be able to perceive the rewards from risky or constant reward flowers as equivalent? Based on the simple model in the text, one-item memory windows are not enough. Consider a memory window of 2 items and the following rule (think like a bee). The bee remembers rewards from two yellow flowers in memory, and averages their values. It stores this value in a ‘comparison circuit’. The two items are used in a computation rule as follows:

$$(\text{reward on 1st} + \text{reward on 2nd})/2.$$

The bee then compares these values to the average reward from a different color of flower. In the later case, the bee must possess another two memory slots that could be used for storing the information on average reward for each color of flower (it already has the average reward from the first flower stored in a comparison circuit). It can then base its decision to forage on one color of flower or the other using these two bits of manipulated information -- the average reward from each floral type. Already there is so much to remember so much data to collect. To make a meaningful decision based on at least two samples from each flower, the bee must sample 4 flowers. Let’s think of a foraging bee on a two flower sampling flight. Landing on a constant reward flower will always yield 2 μl so it is of interest to calculate what kind of returns are expected from the variable flower (e.g., yellow in this case) in all possible foraging outcomes. A bee visiting a full yellow flower (6 μl) and an empty yellow flower (0 μl) gets:

$$\text{average reward} = (6+0)/2 = 3 \mu\text{l}, \text{ which occurs with probability } 1/3 \times 2/3 = 2/9,$$

if it encounters two full flowers in a row:

$$\text{average reward} = (6+6)/2 = 6 \mu\text{l}, \text{ which occurs with probability } 1/3 \times 1/3 = 1/9,$$

two empties in a row yields:

average reward = (0+0)/2 = 0 μl , which occurs with probability:

$$2/3 \times 2/3 = 4/9,$$

or an empty then a full yields:

$$(0+6)/2 = 3 \mu\text{l}, \text{ which occurs with probability}$$

$$2/3 \times 1/3 = 2/9.$$

If a bee remembered the last two flowers, the average reward it perceives is given by:

$$3 \times 2/9 + 6 \times 1/9 + 0 \times 4/9 + 3 \times 2/9 = 18/9 = 2 \mu\text{l}.$$

If a bee remembers the last two items and the bee can compare averages from the colors of flowers, the bee now has the cognitive capacity to perceive the average reward of blue and yellow flowers as being equal. Moreover, the probability of seeing a ‘good reward’ from its foraging on two variable flowers is better than 50:50 (i.e., number of cases where the average of two variable flowers is > 2 μl occurs in 5 out of 9 cases). Only 4 out of 9 cases of two variable flowers yields a zero average reward. Thus, the average forager has a better than even chance of seeing a good reward from variable flowers. It won’t reject the variable reward outright, if it bases its decision on remembering two items from each flower and then comparing the averages of these items between blue and yellow flowers. For the case discussed above, a few quick calculations would show that a memory window of 3 would result in an over designed bee. It need not remember 3 items to have enough information upon which it should base its decision. Though this would give it more information it would take at least 6 foraging flights to collect the data. However, remembering only a single item is not enough (see text).

The Search Image

Another problem that an organism encounters during feeding is finding their prey. Prey have evolved elaborate adaptations to be **cryptic**. By evolving cryptic pigmentation patterns that match their background, many prey avoid detection by their visually oriented predators. A predator's cognitive machinery can be shaped by natural selection to be more efficient at detecting the cryptic prey. Predators are thought to form **search images**, which are mental representation of prey stored in memory. Search images help the predator identify particularly cryptic prey that are camouflaged on substrates such as bark, lichen, or rock. The search image is thought to entail a short-term perceptual change that temporarily increases the predator's ability to detect particularly cryptic prey owing to successful encounters with cryptic prey during past bouts of foraging (Lawrence 1985; Guilford and Dawkins 1987). The formation of a search image is a cognitive process entailing perception of the cryptic prey, data manipulation from a sequence of encounters with a common cryptic prey item, and a representation of the environment forms that allows the predator to reliably detect the prey against the background environment.

Consider the mammalian predator *Barry sinervicus*, a lizard hunter (predator is a euphemism here as *B. sinervicus* does not actually eat lizards, he releases them unharmed). During the breeding season of lizards, *B. sinervicus* catches and marks several thousand lizards in a long-term study of lizard behavior and natural selection. Catching lizards is done with a long fishing pole, which has a noose fashioned from silk thread. The noose is carefully slid over the lizard's neck. However, to "noose" a lizard one must first see the lizards. Lizard back patterns make them quite cryptic against lichen-encrusted rocks. As *B. sinervicus* wanders through lizard land in search of his next quarry, he receives positive reinforcement for every lizard captured (perhaps he feels success, elation, and a sense of a job well done). Negative reinforcement comes in the form of a long waiting time until the next capture (he becomes bored as he stumbles under the hot desert sun not catching any lizards).

What cognitive processes are going on in *B. sinervicus*' head as he tries to form a search image of the cryptic lizard? While I have done no

experiments, *B. sinervicus* may not form an image of the cryptic lizard on the rock as much as he forms an idea of the rock not quite looking right. The distorted rock may turn out to be a lizard on top of a rock. Near the end of the morning of lizard catching he is getting red-hot and he rides the wave of a catching streak. Lizards are rolling in at an average of 20 an hour. At noon he changes over to catching a different species. As he changes to a different species, *B. sinervicus* fails to see any lizards. There is a bit of adjustment time as he makes the perceptual shift from the back pattern of the first species to the back pattern of the next. The time between lizards is long and dry. Why does it take that long to catch the next lizard when it is a different species even though the basic form is similar? Presumably, his cognitive processes require a retooling. Perhaps he needs to rebuild his search image.

This kind of learning process may form a constraint on optimal foraging in animals. If an animal must learn about its environment during foraging, then it is not as efficient as the all-knowing and omniscient forager. While learning gives flexibility to adapt to the changing environment it also brings with it the penalty of decreased efficiency -- at least initially. A classic perceptual constraint arising from the formation of one search image is that the cues that are used to detect the crypsis of one species interfere with the formation of new cues for detecting new cryptic prey items (Tinbergen 1960; Croze 1970). The constraints on perceptual systems result from the data manipulation stage of cognition, in which formation and storage of salient cues derived from past images interferes with the formation of new images.

Perceptual constraints and crypsis

A classic example of crypsis is found in the moth *Biston betularia* from the United Kingdom. In the 1950's Kettlewell reconstructed the historical spread of coal-fired industries in England (Kettlewell 1970). Using information from Victorian-era moth collections (the British upper class developed superb collections of insects), Kettlewell correlated an increase in frequency of a dark versus light morph of *B. betularia* with the advent of the industrial revolution that swept through the United Kingdom. Around areas with high rates of coal burning, the trees became sooty and a melanistic form of moth increased in frequency relative to the light-colored "typical" morph usually found on the grayish lichen in unpolluted areas. Could this change in color affect

a moth's survival (Kettlewell 1955; Kettlewell 1977)? Is there additional natural selection on the cryptic morphs that further perfects their camouflage? Predators can find a mottled gray moth on a black background more readily than they can a black moth on a black background, and this selectivity shaped the frequency of moths on the different substrates.

Kettlewell's natural history observations and subsequent experiments on avian detection systems provide us with the natural history behind a more intricate question regarding the behavioral discrimination abilities of avian predators feeding on cryptic prey items such as moths. The ability of birds to pick out cryptic prey (e.g., gray moths on gray lichen) compared to prey with inappropriate crypsis (e.g., gray lichen on sooty tree bark) should come as no surprise. How do the abilities of the predator change when they are feeding on more than two species, both of which are cryptic, but in a different way?

Pietrewicz and Kamil (1979) tested aspects of the search image hypothesis using **operant conditioning** on blue jays, *Cyanocitta cristata*. Under operant conditioning, the test animal is trained to operate a device such as a button or bar that dispenses a reward. In the case of the blue jay's trained by Pietrewicz and Kamil, the reward was only dispensed if the birds correctly identified a moth that was cryptically concealed on a lichen-encrusted background. In the cryptic prey discrimination test, the blue jays pecked at a key to receive a reward for a correctly identified cryptic moth. Not all the slides that Pietrewicz and Kamil showed the birds contained an image with a cryptically concealed moth. Half of the slides had no moth and if the blue jay pecked at these slides no reward was delivered. The time it took for the next slide to appear in the presentation was lengthened to penalize jays for making an incorrect peck (60 seconds for a wrong answer, and 4 seconds for a correct answer). Presumably birds, like hypothetical herpetologist, *B. sinervicus*, find this delay between potentially rewarding situations unpleasant and it negatively reinforces the wrong choice made in early trials. The blue jay gradually learns to discriminate under the operant conditioning of the "Cryptic Prey Slide Show". Their basic operant conditioning experiment presented two kinds of slides during the blue jay's **discrimination task**:

1. a moth present on a cryptic background,

2. no moth present.

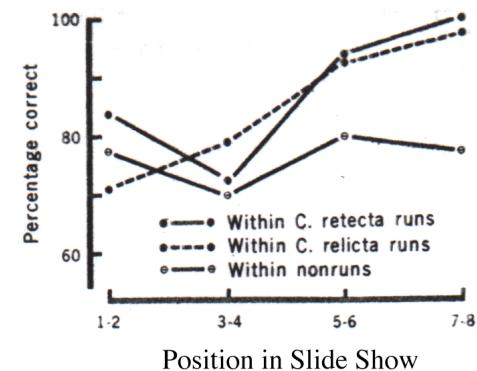
The experiment had the following **reward structure**:

1. If the blue jay pecked at the key when shown a cryptic moth, they received a mealworm larva (*Tenebrio* sp.) as a reward.
2. If they pecked with no moth present, no food was delivered and the jay had to wait before it could see the next slide.

To test the search image concept, two different species of moths with different patterns of crypsis were used, *Catocala relicta* and *C. retacta*. *Catocala relicta* has white forewings with prominent stripes of black and gray and is usually found resting head-up on white birch. *Catocala relicta* has gray forewings with a prominent disruptive pattern of brown lines and is usually found resting head down on dark bark such as oak. Pietrewicz and Kamil only presented pictures of the moths on their suitably cryptic background substrate.

The first set of trials entailed slide presentations of each single species on its species typical substrate. These single-species runs were intermixed with blank slides of the background present. Eight slides with a moth (positive) and without a moth (negative) were presented during one session of 16 slides. This would give the experiment two distinctly different kinds of single-species runs. As a control, Pietrewicz and Kamil ran a 16 slide series with 8 negative slides intermixed with both *C. relicta* and *C. retacta*.

Figure 6.13. Blue jays, *Cyanocitta cristata*, increase in their discrimination of cryptic prey to nearly 100% when they encounter only a single prey type during any one run of a slide show in which empty slides of the background are intermixed with cryptic prey on the background. However, discrimination ability does not improve when two different species of cryptic prey are intermixed with empty background slides in a single run. (from Pietrewicz and Kamil, 1979).



If birds form a search image, then the birds should gradually improve in accuracy over each trial. In addition, if formation of a search image is constrained by cognitive processes blue jays may only be able to form a single search image. Blue jays that are faced with two or more cryptic images, which are mixed together during the same session, would not be expected to learn any search image. Presumably the cognitively challenged blue jay is “maxed out” at a single search image. As expected, the jays rapidly increased to nearly 100% success rate from a background success rate of 80% in the single-species runs. In contrast, the jays did not increase their success rate over the course of trials where the alternative prey was intermixed (Fig. 10.13).

The search image hypothesis has broad intuitive appeal for explaining some of the limits on foraging. It may even extend to all kinds of foraging, not just foraging for highly cryptic prey. It appears to explain how chicks forage for seeds on the ground (Dawkins 1971) or how quail, *Colinus virginianus* (Gendron 1986) and blackbirds, *Turdus merula* (Lawrence 1985; Lawrence 1985; Lawrence 1986) might forage for artificial bait. However, the search image hypothesis as tested in these studies has undergone critical re-appraisal by Guilford and Dawkins (1987). The concept of a search image stipulates that a mental representation of the cryptic prey on its background is formed, or, alternatively some representation of the background with the disruptive form of the prey is formed in the predators brain. This template is then matched with the predator’s current visual field and the prey is detected. Guilford and Dawkins (1987) indicate, quite correctly, that the predator might be paying much more attention to a particular background when a prey item on such a background has successfully rewarded the predator in the past. Indeed, the hypothetical herpetologist *B. sinervicus* might pay more attention to green encrusted lichens because the time between catching lizards is short from such rocks. Thus, a search image might not actually be formed, but rather predators spend more time scanning a background that yields rewards. This leads to different kind of constraint on searching for cryptic prey that is reflected in a speed versus accuracy trade-off, a well-documented psychological phenomenon (Eriksen and Schultz 1978). To detect the subtlety of cryptic prey against the background, more time is necessary to acquire the high acuity information, which is used to base a decision.

This alternative explanation would still imply that the formation of a search image arises from cognitive architecture, but that the constraint arises from the time constraints during the process of perception. This contrasts with the classic search image hypothesis where the constraint is thought to arise from the stages of data manipulation occurring in the brain where the search image is formed from successive prey encounters. The issues of search image are a fascinating example of the interplay between proximate (perceptual and cognitive) and ultimate factors (foraging and fitness), and we will examine the proximate issues at greater depth in Chapter 14. Regardless of whether the search patterns of predators are constrained by data manipulation (storage of successive cryptic images), or by perception (time vs. accuracy tradeoff), such cognitive constraints ultimately constrain foraging patterns of animals.

Summary: Optimal Foraging and Adaptational Hypotheses

The assumption of adaptation is central to optimal foraging theory. The principle of adaptation maintains that the process of natural selection has shaped the behaviors we observe in animals. Successful foragers are favored by natural selection, which results in a species that is imbued with decision rules that maximize energy gain. Gain per unit of time is only one possible currency, and different animals might use alternative currencies. Oystercatchers maximize energy per unit time. Parental starlings maximize energy returned to offspring. Honeybees seem to maximize energetic efficiency that enhances lifespan and utility of the individual worker to the entire colony. Natural selection is assumed to be such an efficient process that animals in nature will invariably be foraging by some optimal rule, or so we are led to believe. In point of fact, natural selection on foraging individuals is often strong (e.g., African seed crackers), so adaptation cannot be perfect. Throughout this chapter, we have embraced the adaptational approach, which has served as a cornerstone of behavioral ecology for more than 40 years.

Gould and Lewontin attacked the adaptational paradigm in 1979. Gould and Lewontin addressed many weaknesses in the adaptationist’s paradigm, but the most difficult challenge related to the notion that all phenotypic traits in an organism, behavioral traits included, must be the result of natural selection. Gould and Lewontin’s arguments related to what they referred to as the “**Adaptationist’s Programme**”. Programme is both a reference to a way of doing science, and a metaphor for a

computer program that would describe how adaptationists carry out research. The basic "adaptationist's programme" described by Gould and Lewontin (1979) would flow as follows:

1. identify the trait that is likely to be under selection,
2. construct a functional argument for how that trait might lead to adaptive value,
3. in the case of optimal foraging theory, assume that energy maximization is related to fitness,
4. determine whether the mean value of the trait in the population is "optimal" in terms of organismal design and function (e.g., energy maximization),
5. if the trait is not found to be under optimizing selection, then go back to step 2 and construct a new functional argument or come up with another trait or perhaps **other factors not considered** that might lead to the discrepancy between the observed value of the trait and the "optimality arguments".

The last step, an iteration back to a previous step, is how the metaphor relates to a computer program. The process used by Meire and Ervynk (1986) to study oystercatchers (see Side Box 6.1) epitomizes the approach that has been used with great success by optimal foraging model builders. Gould and Lewontin (1979) would argue that adaptationists are sure to come up with the right answer and validate their adaptational views because of the recursive nature of their tests. By looping back to the beginning if you do not get the correct "optimal" answer, the adaptationist will eventually stumble upon an adaptational explanation, even if this is not the way the world really works. The adaptationist could stop there, or they could continue looping through the program. The factors considered on each iteration of the loop are:

1. additional energy constraints that influence the efficiency of the forager (e.g., hardness of prey), or
2. additional time constraints arising from the way prey is distributed in the environment.

If they fail to find such an optimal solution, the adaptationist would say 'a mismatch was found between optimal foraging theory and data because of factors and constraints that have not yet been considered'. This caricature of the adaptationist's paradigm presented by Gould and

Lewontin is somewhat unfair. It is hard to get away from the form of hypothesis testing that is used in the adaptationists programme because every science first seeks simple answers to patterns. If the simple answers do not suffice and the models of the simple behavior do not predict the observed behavior, then more complex behavior is incorporated into the model. Even though the characterization of adaptationism by Gould and Lewontin was somewhat reactionary, their attacks did force the entire field to look closely at the inferences people were making regarding the utility of traits under study. Gould and Lewontin suggested that alternative, non-adaptational hypotheses might explain organismal traits. Gould and Lewontin also argued that many of these alternative hypotheses were not being explored during the process of building an optimal foraging model. Many of their arguments relate to constraints on design that arise from organismal architecture.

Behaviorists became more sensitized to alternative non-adaptational hypotheses. The development of cognitive models of foraging grew from the need to build in the constraints of neural architecture into the decisions that are made by animals. Cognitive constraints of memory and perceptual constraints entailed in search images may model how a nervous system is built, but it is likely an oversimplification. In upcoming Chapters we will explore additional constraints of architecture on the behavior of animals. The cognitive approach to behavior can enrich our understanding of the biological bases of decision rules. Much work on the neural basis of decisions remains to be done.

Optimal foraging is a mature and rich field, but there are still issues to address. The assumption in the third step of the adaptationist's programme is critical to optimal foraging theory. We assume that energy maximization or optimization is in some way related to fitness maximization (e.g., stabilizing selection on the trait of interest). Much of optimal foraging theory rests on this assumption and a few studies have successfully demonstrated the link between foraging and fitness in nature. The optimal foraging demonstrated by foraging seed crackers provides us with an example of strong links between foraging efficiency of small- and large-billed birds, and survival of these birds in the wild. In the same vein, the developments of risk sensitive foraging place animals in much more realistic experimental contexts that are likely to be faced in the wild. Animals in nature forage in the face of tremendous

uncertainty. Laboratory experiments shows that most animals tend to avoid risk. When food is available, there are no gains to be had from feeding on risky food sources. However, when starvation probability is high, animals become gamblers and they forage on food with a variable payoff in the hopes of striking it rich and weathering through the hard times. Nature selects for a certain amount of gambling in the decision making of animals. Applying these principles in the wild is the next step, although experiments already indicate that wild bees appear to be risk sensitive as a function of their colony's reserves.

Study Questions for Optimal Foraging

1. Describe the central assumption underlying optimal foraging theory. How would a student of the adaptationists programme modify their assumptions if they did not find that an animal was foraging optimally?
2. Describe the problem faced by a central place foraging chipmunk. What additional energy considerations does a parental starling face from the viewpoint of central place foraging?
3. Describe how learning as a cognitive process, and memory as a cognitive process limit the solutions that animals can come up with from the point of view of optimal foraging? Do these processes constrain optimal foraging and if so, why?
4. Why might animals appear to risk aversive? Explain in terms of a bumblebee foraging on flowers and a cognitive model of memory.
5. Why are animals gamblers? What is the adaptive benefit of this behavior? Give an example of an animal likely to be risk prone.
6. Apply the “adaptationist’s programme” to optimal foraging in oystercatchers. List the factors that you would have to build into your optimal foraging model.
7. Describe three different currencies that might be optimized in animals (hint also consider levels of selection) (Chapter 4).