LUND UNIVERSITY

The foraging benefits of information and the penalty of ignorance

Olsson, Ola; Brown, JS

Link to publication

Total number of authors:
2

The foraging benefits of information and the penalty of ignorance

Ola Olsson and Joel S. Brown

Olsson, O. and Brown, J. S. 2006. The foraging benefits of information and the penalty of ignorance. – Oikos 112: 260–273.

Patch use theory and the marginal value theorem predict that a foraging patch should be abandoned when the costs and benefits of foraging in the patch are equal. This has generally been interpreted as all patches being abandoned when their instantaneous intake rate equals the foraging costs. Bayesian foraging – patch departure is based on a prior estimate of patch qualities and sampling information from the current patch – predicts that instantaneous quitting harvest rates sometimes are not constant across patches but increase with search time in the patch. That is, correct Bayesian foraging theory has appeared incompatible with the widely accepted cost–benefit theories of foraging. In this paper we reconcile Bayesian foraging with cost–benefit theories. The general solution is that a patch should be left not when instantaneous quitting harvest rate reaches a constant level, but when potential quitting harvest rate does. That is, the forager should base its decision on the value now and in the future until the patch is left. We define the difference between potential and instantaneous quitting harvest rates as the foraging benefit of information, FBI. For clumped prey the FBI is positive, and by including this additional benefit of patch harvest the forager is able to reduce its penalty of ignorance.

O. Olsson and J. S. Brown, Dept of Biological Sciences, Univ. of Illinois at Chicago, 845 W. Taylor St, 60607 IL, USA. Present address for OO: Dept of Animal Ecology, Lund Univ., Ecology Building, SE-223 62 Lund, Sweden (ola.olsson@zooekol.lu.se).

A forager's food may be distributed patchily. Spatial variability in food availability poses opportunities and challenges to the forager. As opportunity, variability allows the forager to bias its searching efforts towards rich patches and away from poor patches (Stephens 1989). Most models of patch use (Charnov 1976, Oaten 1977, Brown 1988) and habitat selection (Fretwell and Lucas 1970, Rosenzweig 1981) assume that foragers assess spatial heterogeneity in feeding or fitness opportunities and respond accordingly. As challenge, though, the forager must be able to assess this heterogeneity before it can benefit from a more efficient allocation of effort. Furthermore, variability in food availability can lead to variability in the individual's food intake rate. Depending on the relationship between food consumption and fitness (often an increasing and decelerating curve) a forager's fitness may be influenced by temporal variability in feeding rates (for instance, risk sensitive foraging, Caraco 1980, Real 1980). Here we are interested in the assessment challenge posed by patchily distributed foods.

Forager's may obtain information from a variety of sources. Prior to selecting a food patch, a forager may gain knowledge on the whereabouts of rich and poor patches from visual, auditory or chemical cues that can be detected at a distance. Such long distance cues include patch appearance (e.g. floral color or abundance, Sandlin 2000; the fruit scent or burdens of trees, Sallabanks 1993) or observations on the foraging successes of other individuals (Valone and Giraldeau 1993). This allows a forager to be "periscopic" (sensu Mitchell 1989) by selectively visiting higher quality patches and/or by optimizing the order within which patches are visited (variations on the "traveling salesman problem", Gross et al. 1995). Upon encountering a patch, a forager may gain information on its quality,

usually represented as the abundance of resources. A forager may be prescient (sensu Valone and Brown 1989). Upon entering the patch, such a forager uses cues from the patch to make the most accurate assessment of quality. Alternatively, the forager uses information gained while exploiting the patch to make and continuously update an estimate of patch quality (Oaten 1977, Green 1980, Iwasa et al. 1981, McNamara 1982, Olsson and Holmgren 1998). Here we are interested in this last information challenge. Specifically, how should a forager use knowledge of cumulative search time and cumulative harvest within a patch to effect an optimal patch use decision?

Bayesian foraging has provided the conceptual framework for how a forager should estimate patch quality from three sources of information: 1) an a priori knowledge of the distribution of patch qualities throughout the environment, 2) time spent thus far in searching for food items within the current patch, and 3) number of food items thus far encountered and harvested. When search within a patch is random (equal and constant encounter probability on all food items within the patch), authors agree on how a Bayesian forager can estimate current patch quality.

However, authors disagree on how a fitness maximizing forager should use this information to decide how thoroughly to use each food patch. Cost–benefit models of patch use suggest that a forager should remain in the patch until the expected instantaneous harvest rate no longer exceeds foraging costs (Charnov 1976, Iwasa et al. 1981, Valone and Brown 1989). If foraging costs do not vary among patches, then the forager should strive to equalize quitting harvest rates among patches (Brown and Mitchell 1989) and leave each patch when its estimated quality falls to a threshold giving-up density (GUD, Brown 1988).

While seductively straightforward, the balancing of expected instantaneous harvest rate with foraging costs can be wrong for a Bayesian forager (Green 1980, 1984). That is, such a forager does not necessarily maximize its fitness by leaving all patches at the same instantaneous intake rate (Olsson and Holmgren 1998, 2000). Here we are interested in reconciling the attractiveness of cost–benefit models of patch use with the following conceptual facts about Bayesian foraging and patch use.

The peculiarities of Bayesian patch use and its incompatibility with traditional patch use models such as the marginal value theorem (Charnov 1976), or $H = C + P + MOC$ (Brown 1988) go back to Oaten (1977), Green (1980) and McNamara (1982). But, it is a technical report by Green (1988) that fully illuminates the potential incompatibilities of a fixed quitting harvest rate patch use rule with optimal foraging under Bayesian patch use. Iwasa et al. (1981, and Valone and Brown 1989, Rodríguez-Gironés and Vásquez 1997) merely assume a fixed quitting harvest rate strategy as optimal. Green shows beautifully how this is true only when resources are not clumped (e.g. binomial or Poisson distributions).

When the distribution of resources among patches is clumped (e.g. negative binomial – most patches have few items and a few patches have many items), a forager should use a potential value rule and only leave the patch when the expected average harvest rate within the patch for the remainder of the visit no longer exceeds foraging costs. Olsson and Holmgren (1998) showed how this subtle yet crucial distinction leads to an interesting pattern of patch use. Instead of a constant relationship between quitting harvest rate and time spent within a patch, quitting harvest rate actually rises with patch residence time. They show how a forager may be willing to persevere in a seemingly unsatisfactory patch in the knowledge that finding a food item may provide the good news that this patch is not so bad after all. Here, we show how incorporating this "good news" into patch use models can reconcile potential value rule of Green (1988) and Olsson and Holmgren (1998) with the instantaneous rate rule of the patch use model of Brown (1988, 1992) and with the marginal value theorem (Charnov 1976).

Our specific goals include:

1) Highlight the salient differences and consequences of a Bayesian forager using a fixed quitting harvest rate strategy versus a potential value rule. Under clumped distributions of food, the fixed quitting harvest rate strategy that seems so compatible with cost–benefit models of patch use is wrong. The potential value rule, which is optimal, does not seem compatible with traditional cost–benefit models.

2) Introduce the concept of a foraging benefit of information (FBI). With this concept we suggest that the optimal patch use strategy of a Bayesian forager should be: $H + FBI = C + P + MOC$ where H is the expected instantaneous harvest rate, and C, P, and MOC are the metabolic, predation and missed opportunity costs of foraging, respectively.

3) By using a prescient forager as a benchmark we clarify several aspects of patch use under imperfect information. First, there is always a penalty of ignorance when a forager is not prescient. Next, from all available information, useful information is that which can be used to reduce the penalty of ignorance. Finally, the useful information from patch exploitation that can lead to good news and raise the forager's estimate of current patch quality leads to a positive FBI.

Models and concepts

We illustrate the results by considering two ecological scenarios that differ in their foraging costs. The first scenario conforms to the assumptions of the marginal value theorem. The environment is infinitely repeating meaning that while there is depletion of harvested food patches, the environment itself is not depleting. The forager's objective is to maximize its long-term net energy gain in the absence of predation or metabolic costs of foraging. The only foraging cost while harvesting a patch is the missed opportunity from not leaving the patch and exploiting another. This missed opportunity cost of the marginal value theorem is just the forager's long-term average intake rate, Γ. This is the scenario discussed above where each patch use strategy involves leaving a patch when the rule's value function falls to the threshold of Γ.

In the second scenario we relax the assumptions of the marginal value theorem by assuming that there are both metabolic and missed opportunity costs that determine the threshold at which the forager should leave the food patch. We assume that these costs are exogenous to the distribution of food items among patches and to the forager's patch use strategy. In this scenario, the forager's goal is to remain in each patch until the value function equals its exogenous cost. Unlike the first scenario, each rule under this second scenario can have the same exogenous cost. And, each rule may strive to harvest the same amount of food from each patch. A rule outperforms another rule by more successfully biasing search effort towards rich patches and away from poor patches.

For each scenario, we will consider the performance of each patch use rule for a range of distributions of food items among patches. Using a binomial distribution we will consider the range from completely uniform to Poisson, and using the negative binomial we can continue the range from Poisson to clumped.

The analyses will reveal the penalty of ignorance, and show the failure of the Bayesian instantaneous rule relative to the Bayesian potential value rule. The analyses will allow us to formalize a benefit from spending time searching a patch. We will term this benefit the foraging benefit of information (FBI). We will see how the Bayesian instantaneous rate rule fails because it either ignores or fails to properly incorporate FBI, and how the Bayesian potential value rule succeeds by correctly incorporating the FBI into the costs and benefits of foraging.

Scenario 1: intake rate maximization

We will compare the results of four different patch-leaving rules. To begin our analyses, we assume that the forager has no metabolic or predation costs (or these

costs are the same for all activities). Hence, the foragers strive to maximize the long-term rate of energy gain:

$$\Gamma = \frac{\bar{n}}{\bar{t} + \tau} \qquad (1)$$

where $\bar{n}$ is the average number of prey caught per patch, $\bar{t}$ is the average search time spent per patch, and τ is the average travel time between patches.

Within patches we consider foragers whose patch harvest rate, f, conforms to a random search model such as Holling's disc equation:

$$f = \frac{AX}{1 + AhX} \qquad (2)$$

where A is the forager's encounter probability on food items in the patch (or searching efficiency), h is the handling time per harvested food item, and X is the current food density within the patch. Now, let patches vary in their initial prey abundance, N_i, where $i = 1, \ldots, m$ represents m different initial abundances among patches. How does information regarding the initial abundance of patches enhance the forager's performance?

Here, we will consider three general distributions that may be applicable to patches containing a discrete number of prey items: the binomial distribution, the Poisson distribution and the negative binomial distribution. These distributions form a continuum from regular to clumped. For simplicity, we will describe all three distributions with their common parameters $\overline{N}$ and s^2, i.e. the mean number and variance of prey items in patches.

For ease of presentation, we will assume that $h = 0$. This means that rates and prey density match perfectly and can be compared on the same scales. That is, Eq. 1 becomes $f = AX$.

The models of patch assessment that we consider are:

1) Prescient (PS). Upon arrival in the patch the forager immediately has an accurate "knowledge" of patch quality. The forager's value function for a given patch is its current expected harvest rate:

$$f = A(N - n) \qquad (3)$$

where N is the initial number of food items in the patch and n is the number of food items harvested thus far. The challenge for the prescient forager is to select k, the threshold remaining food density that maximizes long-term harvest Γ. If the patch contains less than k prey at arrival the patch should be left without spending any time there. Otherwise, the patch should be harvested until only k prey are left. The leaving rule that accomplishes this is to choose k* such that the value function equals the long term average: $Ak^* = \Gamma$ which can be rearranged as

$$k^* = \frac{\Gamma}{A} \qquad (4)$$

Because the forager leaves the patch when its quitting harvest rate equals its long-term average, this model represents the discrete version of the marginal value theorem (Charnov 1976) presented by Olsson and Holmgren (1999).

2) Fixed time (FT). The forager makes no assessment of the quality of a given patch. Because it cannot discriminate among patches, its optimal behavior is to leave all patches at the same $t^*(=\bar{t})$. The fixed amount of search time should be selected so that the forager leaves each patch at an expected quitting harvest rate that equals its long-term average. The value function of a patch is its current expected harvest rate. Recall that under random search the number of prey caught is

$$n = N(1 - e^{-At}) \qquad (5)$$

(Olsson and Holmgren 1999), and hence the prey density remaining is

$$X = Ne^{-At} \qquad (6)$$

The value function is then

$$A\overline{N}e^{-At} \qquad (7)$$

where $\overline{N}$ is the average initial value of a resource patch. The optimal fixed search time, t*, satisfies $A\overline{N}e^{-At^*} = \Gamma$. By substituting Eq. 1 for Γ and Eq. 5 for n, this expression for t* can be rearranged and given as:

$$e^{At^*} = 1 + A(t^* + \tau) \qquad (8)$$

Notice how the optimal amount of time to spend searching each patch is independent of the average quality of patches and the distribution of resource among patches. It is fully determined by travel time and the forager's encounter probability on resources. Increasing travel time increases t* while increasing the encounter probability decreases t*. Spending t* per patch is the very best that an ignorant forager can do, and such a forager need only know its encounter probability on resources and its travel time among patches.

3) Bayesian instantaneous rate rule (BI). This forager knows the amount of time that it has spent searching the patch, t, the number of prey items encountered thus far, n, and the distribution and abundance of prey items among patches. Such a forager can use this information to estimate the current prey density of the patch. Equivalently, it is capable of determining its expected instantaneous harvest rate. Under this patch departure rule a patch is abandoned when the expected instantaneous intake rate, r, falls to some constant value, Q. Iwasa et al. (1981) showed that

$$r_{n,t} = \frac{\lambda + n}{e^{At}\dfrac{\alpha + 1}{\alpha} - 1} \qquad (9)$$

This is the forager's value function and gives the expected instantaneous harvest rate from remaining in the patch. In Eq. 9, λ and α are parameters from the negative binomial distribution, such that $\lambda\alpha$ is the mean, and $\lambda\alpha(1+\alpha)$ is the variance. By substituting this into Eq. 9 we get:

$$r_{n,t} = \frac{n(s^2 - \overline{N}) + \overline{N}^2}{s^2(e^{At} - 1) + \overline{N}} \qquad (10)$$

Equation 10 is valid for all three distributions considered here (binomial, Poisson and negative binomial). As can be seen from this expression, the behavior of r varies with prey distribution (Iwasa et al. 1981). In a binomial distribution $s^2 < \overline{N}$ and hence r decreases with n. In a Poisson distribution, $s^2 = \overline{N}$ which means that r becomes independent of n. In a negative binomial distribution, $s^2 > \overline{N}$ and r increases with n. In all three distributions, r declines with t. However, in a completely uniform distribution, where $s^2 = 0$, r becomes independent of t but declines with n. Iwasa et al. (1981) proposed that the forager should leave all patches when r drops to some constant level Q*. They will then achieve some long-term intake rate, Γ.

4) Bayesian potential value rule (BP). This rule considers a forager with the same information state as the previous rule. From the information on n and t the forager can, in addition to estimating the expected instantaneous intake rate, look ahead and anticipate its prospects from harvesting the patch for an additional finite amount of time. By anticipating the future of harvesting the patch, the forager can estimate the expected number of prey that will be caught during the remainder of the patch visit, E(n), and the expected amount of time that will be spent searching for those prey E(t). The ratio of these provides the forager's value function. This value function estimates the forager's average return from the patch during the remainder of its patch visit, Π. This ratio

$$\Pi = \frac{E(n)}{E(t)} \qquad (11)$$

can be called the "potential intake rate" of the patch (Olsson and Holmgren 1998). Under the potential value assessment rule it is optimal to leave all patches when Π reaches some constant value, Π^* (Green 1988, Olsson and Holmgren 1998). For a rate maximizer this value should be $\Pi^* = \Gamma$ (Green 1988).

The long-term energy intake rates that can be achieved using the four different models will be called Γ_{PS}, Γ_{FT}, Γ_{BI}, and Γ_{BP} respectively (PS =prescient, FT =fixed search time, BI =Bayesian instantaneous rate rule, and BP =Bayesian potential value rule).

Scenario 2: fixed exogenous cost

For generality, we will also consider foragers that have metabolic costs of foraging, deplete their environment, and have the option of saving energy by remaining inactive. We will let c be the metabolic cost of foraging and c_0 be the cost of resting in a refuge or burrow. For simplicity we will not consider any predation risk in this model.

In this model, the forager has to pay c_0 for each time unit whenever it rests and does not forage. Hence, this forager's missed opportunity cost is $-c_0$, and its quitting harvest rate should therefore be H =c $-c_0$ (Brown 1992).

The rate maximizers' missed opportunity cost, and hence quitting harvest rate, varies widely with proficiency between the different patch use rules. In this case quitting harvest rate is fixed and the same for all patch use rules, as they are dictated by c and c_0 which are the same regardless of strategy. This case therefore illuminates several important properties of the patch use rules, which will be evident below.

We assume that this forager has a time horizon that is T time units, and its environment (e.g. territory) consists of K patches. It has three possible activities: foraging in patches, traveling between patches or resting. On average it will gain $\bar{n}$ food items (each containing one unit of energy), and spend $\bar{t}$ time units per patch foraging and spend τ time units traveling between each patch. As mentioned above, when foraging or traveling it will spend c energy units per time unit, and when resting it will spend c_0.

The net energy intake rate of a forager like this is hence:

$$\Gamma = \frac{K\bar{n} - cK(\bar{t} + \tau) - c_0(T - K(\bar{t} - \tau))}{T} \quad (12)$$

and this is the fitness function it strives to maximize. Equation 12 is subject to the constraint that there has to be sufficient time within T such that it can deplete all the K patches down to its desired level. That is $K(\bar{t} - \tau) < T$ must be satisfied.

For each of the patch use rules we choose the policy that maximizes Γ. In the case of the fixed-time strategy, the optimal solution can be found analytically. As we assume random search, we may substitute Eq. 5 for n into Eq. 12, and differentiate with respect to t. By setting the derivate to zero and solving for t we then find that

$$t^* = -\frac{\ln\dfrac{c - c_0}{A\overline{N}}}{A} \quad (13)$$

For the other patch use rules, the optima must be found numerically.

Penalty of ignorance

The prescient forager's performance provides a ceiling on performance for comparing the performance of the other three patch use rules. For comparison, the fixed time rule of the ignorant forager provides a floor on performance. Between the upper and lower benchmarks of the prescient rule and the fixed time rule, respectively, lie the performances of Bayesian assessment strategies.

We will define the penalty of ignorance, I, as the difference in performance between the prescient rule and the fixed time rule (Fig. 1):

$$I = \Gamma_{PS} - \Gamma_{FT} \quad (14)$$

This difference may be regarded as the selection gradient for improving the information processing capacity of the forager over evolutionary time. A Bayesian forager can only recoup some but not all of the penalty of ignorance by using sampling information.

We will define the value of sample information as the improvement in performance of the Bayesian forager
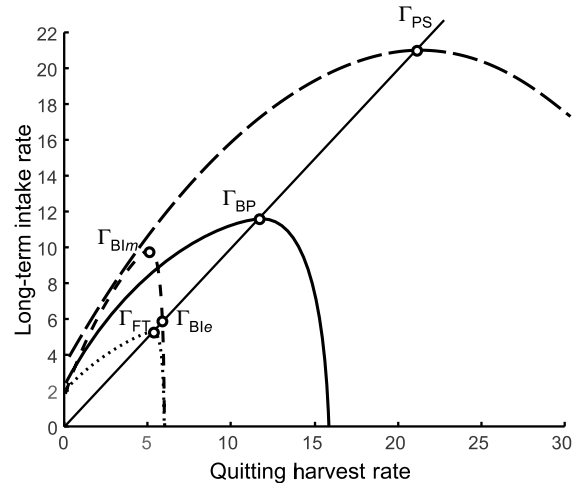

Fig. 1. The long-term energy intake rates returned by the four different models of scenario 1. Open circles indicate the respective maxima. The highest dashed curve is the one produced by a prescient forager, and is shown as a function of k, the quitting harvest rate chosen. The solid curve is for the Bayesian potential value assessment rule, shown as a function of Π, the estimated potential value of the patch at which it is abandoned. The lower dashed curve is the Bayesian instantaneous value assessment rule, shown as a function of Q, the estimated instantaneous value of the patch. Finally, the dotted curve is for a fixed time forager. It is shown as a function of the GUD produced by the strategy t chosen.

over that of the ignorant forager: $\Gamma_{BP} - \Gamma_{FT}$. We will define the penalty of sampling information as that portion of the penalty of ignorance that cannot be recouped through the use of sampling information alone: $\Gamma_{PS} - \Gamma_{BP}$. By definition the value of sampling information and the penalty of sampling information sum to the penalty of ignorance. The penalty of sampling information represents selection for the evolution of other sensory modalities for assessing patch quality over and above the Bayesian use of sampling information.

Results

Rate maximizing performance of patch use rules: scenario 1

For illustration, consider an environment where the prey density among patches fits a negative binomial distribution with a mean prey density of $\overline{N} = 6$ and an overdispersion coefficient of $\lambda = 0.5$ (yielding a variance of 78). Now, consider a patch within this environment with an initial prey density of $N = 24$. In Fig. 2 we show the four patch use rules in the state space of instantaneous intake rate versus time spent searching for food in the patch. The solid stepped line is the actual current prey density. Prey items were found at the times when this line jumps down.

A prescient forager leaves the patch at point A, when the number of prey left reaches $k^* = 21$ and it has caught just three prey items. As can be seen from Fig. 1, the optimum for the prescient forager occurs where $k^* = \Gamma_{PS}$, i.e. when the current value of the patch has reached the best possible long-term rate achievable in the environment. In this example, a prescient forager rejects all those patches with $N \leq 21$ and forages all other patches to $N = 21$.

A fixed time forager leaves the patch at point B after time t^* where the expected instantaneous intake rate, or equivalently average GUD, produced by that t, equals Γ_{FT} (Fig. 1). Note that this decision is independent of the intake rate at that point. In fact this forager, spends this amount of time in each patch irrespective of initial prey abundance and in general harvests ca 1/8th of the food from each patch. In this very example, it takes three items before it leaves, just like the prescient. However, the strategy contrasts sharply with the prescient forager that harvests nothing from patches below $N = 21$ and harvests an increasing fraction of a patch's food as N increases above this threshold.

A Bayesian forager's estimate of current prey density, or equivalently instantaneous intake rate (the solid, jagged curve, in Fig. 2) uses information on the cumulative search time in the patch and the cumulative number of food items encountered. Upon arriving at a new patch a Bayesian forager estimates that the patch
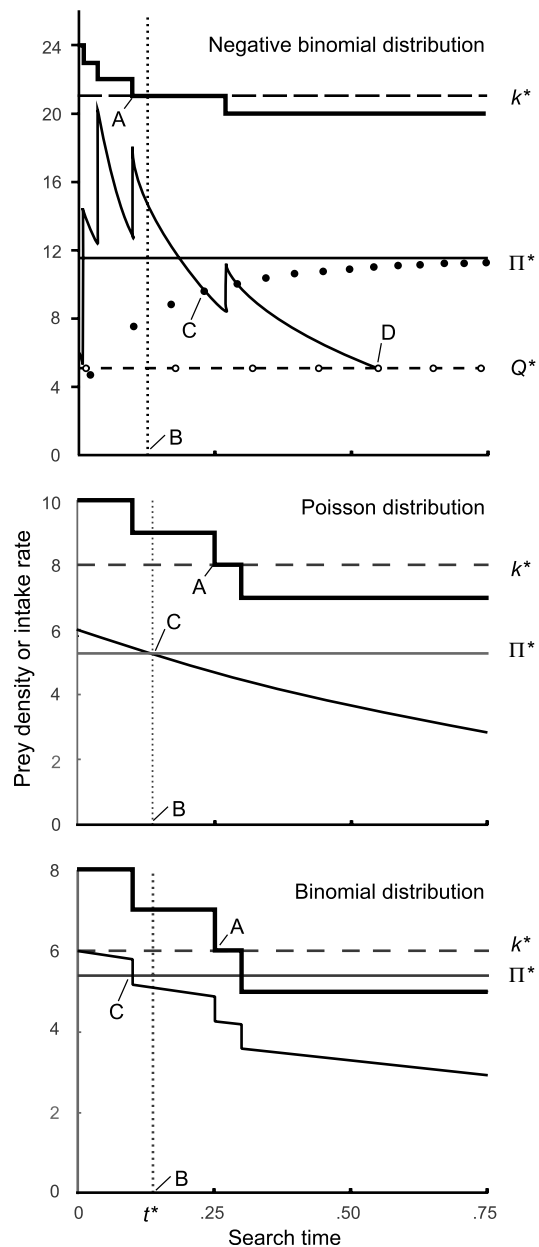

Fig. 2. Actual current prey density (thick step shaped solid curve), estimated current prey density (thinner solid curve), and patch departure rules for four foraging models, exploiting prey of three different distributions. The dashed straight line, at level k*, is the prey density at which a prescient forager should leave. The vertical dotted line, at t*, is the fixed time that a forager unaware of, and incapable of estimating, patch quality should spend. The horizontal solid line, at Π*, is the potential patch quality at which a Bayesian potential value assessment forager should leave. The solid circles are the estimated current prey density at the leaving points corresponding to having found 0, 1, 2 . . . and so on prey items, for this strategy. The dashed horizontal line, at Q*, is the estimated quitting harvest rate of a Bayesian instantaneous value assessment forager in the negative binomial distribution. The open circles are the stopping points for this forager. The letters A–D indicate the optimal stopping points for the prescient, fixed-time, Bayesian potential value, and Bayesian instantaneous value foragers respectively.

contains the average prey density of N = 6. This estimate deteriorates as search time is spent without finding prey. However, when a prey item is found the estimate jumps up; very much so when prey is found soon, and less so when prey are found after longer search times. Hence, we make the mildly paradoxical observation that finding an item leads to an increase in the estimate of current prey density even as the actual prey density has declined.

Following the Bayesian potential value assessment rule, the forager leaves the patch having caught three items at point C. The decision to leave is based on the potential value of the patch, Π, that has reached $\Pi^* = \Gamma_{BP}$ at this point (Fig. 1). It is worth noting that the instantaneous intake rate (or current prey density), indicated by the filled circles, does not have a constant value at a constant Π-value (Fig. 2). Rather, current prey density, or equivalently GUD, on the average across all patches increases with search time (Olsson and Holmgren 1998, 1999, 2000). After long search times, the instantaneous intake rate at departure (H) converges with Π^*.

Finally, a forager following the Bayesian instantaneous value assessment rule leaves the patch when the instantaneous intake rate has reached Q*, at point D. At this point, the long-term intake rate is maximized, but here $Q^* < \Gamma_{BIm}$ (the index refers to Bayesian instantaneous maximizing; Fig. 1)! An alternative would be to leave at a slightly higher instantaneous intake rate, indicated by Γ_{BIe} (Bayesian instantaneous equalizing) in Fig. 1, where instantaneous and long-term intake rates are equalized. However, at this point long-term intake rate is lower! That is, following the Bayesian instantaneous value assessment rule, it does not seem possible to follow the marginal value rule!

The former case corresponds to a naïve MVT in which the forager sets its quitting harvest rate equal to its long-term average. This is what Iwasa et al. (1981) and Valone and Brown (1989) had in mind in their applications of Bayesian foraging. Curiously, this is not always the value of Q that maximizes Γ. When resources are clumped, the optimal value for Q under this patch use rule is to leave patches at a lower quitting harvest rate than the long-term average: $Q^* < \Gamma$. That means that, given the information state of the forager, an instantaneous assessment rule is not optimal when the distribution of prey items among patches is clumped (negative binomial)!

For comparison, consider the cases with Poisson and binomial distributions in Fig. 2. In the Poisson case, the patch initially contains 10 prey items. As a Bayesian forager gains no information from finding prey, its value function is independent of prey captures and it leaves after a fixed time (C), at the same point as the fixed time forager (B). In the binomial case (here with $s^2 = 2$) the patch initially contains eigth prey items. The value function declines with time, but also stepwise as prey are found.

The foraging benefit of information

The Bayesian instantaneous value forager does not achieve the same intake rate as the Bayesian potential value forager (Fig. 1). As mentioned above, this depends on the fact that this strategy fails to incorporate the FBI.

The prescient forager bases its patch departure on the number of prey items left, k, and Ak is the instantaneous quitting harvest rate. Obviously, k also contains all the information there is about the future expectations for the patch – it is the remaining number of prey items, and therefore also the potential value of the patch, and Ak is the potential quitting harvest rate Π.

For prey distributions where $s^2 \leq \overline{N}$ the expected instantaneous intake rate, $r_{n,t}$, can never increase during a patch visit. This means that the expectations for the future during the patch visit are sufficiently described by $r_{n,t}$ and hence FBI = 0. Therefore, $\Pi = H$ in these cases, just like in the case of the prescient forager.

However, as soon as $s^2 > \overline{N} r_{n,t}$ increases with n, and may in the next instant have a higher value than it has now. Thus, $r_{n,t}$ is not a sufficient estimate of the future expectations in the patch. In clumped prey distributions there may be a benefit to remaining even when $r_{n,t} < \Pi$. The benefit from lingering in the patch is the possibility of obtaining positive information about patch quality. The "good news" increases future expectations of prey capture. This benefit of exploiting the patch for information can be measured in units of energy. We define the foraging benefit of information as FBI = $\Pi - H$, i.e. the difference between the potential and instantaneous harvest rates at which a patch is left.

For the Bayesian potential value forager, FBI declines throughout the patch visit. This means that upon arrival the forager accepts a reduced instantaneous intake rate in order to be able to separate rich patches from poor. That is, it gives itself time to make sure that the patch is not much better than it presently thinks – the forager waits for good news (Olsson and Holmgren 1998). When it has spent a long time in the patch FBI is much smaller, as then the estimate is becoming reliable, and even good news change the expectations little.

For a given prey density distribution, FBI is not fixed, as it depends on Π, the quitting harvest rate (Fig. 3A). For a rate maximizing forager only one value of Π can be optimal (Green 1988), but for foragers with other fitness functions other Π-values will be optimal (below, Olsson and Holmgren 1999, 2000). For state dependent foragers a range of different Π-values may be optimal under different conditions (Olsson and Holmgren 1998).

In Fig. 3B we have shown how FBI varies with Π over a large range of values. For comparison the long-term intake rate is also shown in the same graph, but this does not imply that the data are only valid for rate maximiz-
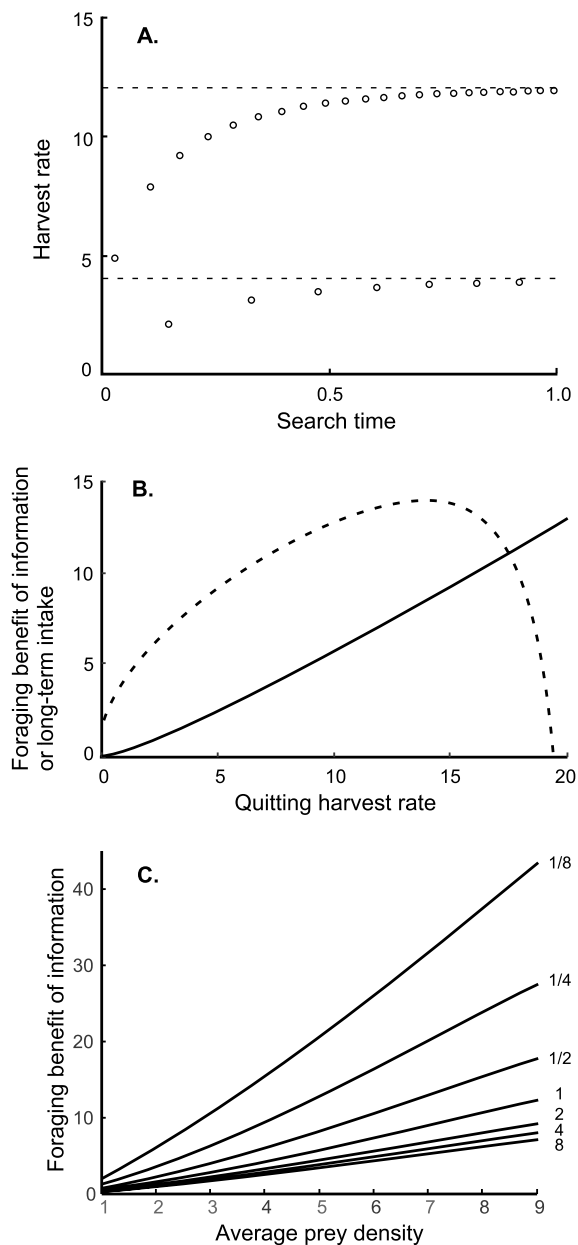
Fig. 3. (A) The potential (Π, dashed curve) and instantaneous (H, circles) quitting harvest rates for two Bayesian potential value foragers that aim to leave the patches at $\Pi = 4$ and $\Pi = 12$, respectively. Both foragers are exploiting the same negative binomial prey density distribution with $\overline{N} = 7$ and $\lambda = 0.5$. The distance between the curves and the first point for each is FBI_0. (B) FBI_0 (solid curve) and long-term intake rate (dashed curve), for the same prey density distribution, for the full range of possible potential quitting harvest rates, Π. (C) FBI_0, for rate maximizing Bayesian foragers, as a function of average prey density. The different curves are for different λ-values, and $s^2 = \overline{N} + \overline{N}^2/\lambda$.

ing foragers. The FBI-value shown is that for the first stopping point (Fig. 3A), when no prey has been found. We may call this FBI_0. It is clear that FBI_0 increases

almost linearly with Π, for a given prey density, and can be rather large in comparison to the longterm gain rate. That is, a large sacrifice in terms of instantaneous intake rate is made to gain information. This information is used to exploit the environment in an optimal manner.

The information available differs much between prey distributions. Information is obviously more valuable in a distribution with high mean density (Fig. 3C). However, it is important to realize that the information available depends very much on the aggregation of food items among patches. In a highly aggregated distribution, with a high variance in relation to the mean, FBI_0 is very large (Fig. 3C). In a less aggregated distribution it is less, and in a Poisson or binomial distribution it is absent! That is, for these two distributions, $\Pi = H + 0$.

Penalty of ignorance

Long-term intake rate increases with the average prey density for all strategies (Fig. 4), but at different rates for the different strategies. This translates into the penalty of ignorance and the value of sample information. The penalty of ignorance declines, but only very slowly with increasing prey density. That is, the richer the environment, the less costly it is to be uninformed. In the example given, the over dispersion parameter $\lambda = 0.5$, and therefore the variance is $s^2 = \overline{N} + \overline{N}^2/\lambda$. That is, the distributions are rather clumped, but the degree of "clumpedness" remains constant.

As opposed to the other strategies the long-term intake rate of the Bayesian potential forager is not linear, but accelerates slightly with prey density. As a consequence, the value of sample information increases with prey density. That is, the relative value of correctly using sampling information increases with environmental quality.

In Fig. 5, we show the results of changing the variance, or the type, of the prey distribution rather than its mean value. Variance ranges from 0 to approximately 300, but the mean value is kept constant at 6 prey items per patch. When $s^2 = 0$ the forager faces a completely regular prey distribution, with 6 prey items in each patch. When $0 < s^2 < 6$ the distribution is binomial, with a symmetrical or left skewed distribution of patch qualities around the mean value. When $s^2 = 6$ it is a Poisson distribution, and when $s^2 > 6$ the forager faces negative binomial distributions, where prey are few or absent in most patches, but some patches are very rich, i.e. a right skewed distribution.

As was shown in Fig. 3 the information content varies considerably with the parameter λ of the negative binomial distribution, and hence with variance (a low λ means a high variance, for a given average prey density). In the range where variance is greater than the mean, the Bayesian potential value forager exploits FBI to increase long-term intake rate (Fig. 5). The outcome is that the value of sample information
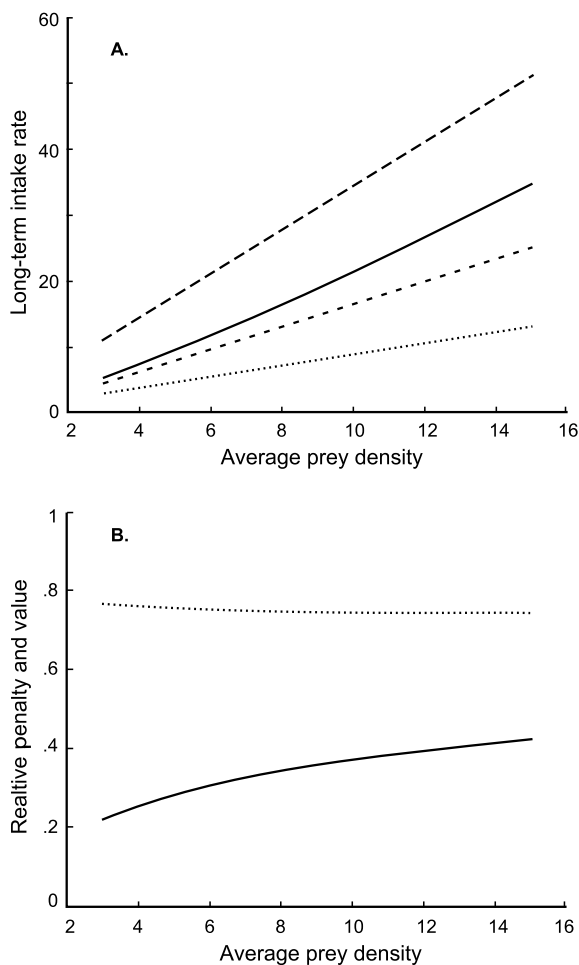
Fig. 4. (A) Long-term intake rate for the four different patch use strategies when mean prey density varies. The upper dashed curve is Γ_{PS}, the solid curve is Γ_{BP}, the lower dashed curve is Γ_{BI}, and the dotted curve is Γ_{FT}. (B) The relative penalty of ignorance, $(\Gamma_{PS} - \Gamma_{FT})/\Gamma_{PS}$ (dotted curve) and value of sample information, $(\Gamma_{BP} - \Gamma_{FT})/\Gamma_{PS}$ (solid curve). In all cases $\lambda = 0.5$.

increases with variance, as long as $s^2 > \overline{N}$. Also the penalty of ignorance increases dramatically with variance. Both increases have the same reason: the more extremely clumped a distribution is, the more readily can bad patches be discarded and foraging effort be allocated to the richest patches.

The fixed time forager has no capability of exploiting the variance of the prey distribution, only its mean. Therefore, its long-term intake rate is independent of variance.

When the variance is not greater than the mean the two Bayesian strategies give identical payoffs. This is because in the binomial and Poisson distributions FBI = 0, and there is no good news to be had from finding a prey item. What may at first seem a paradox is that the intake rate of the Bayesian foragers, and hence the value of sample information, decreases with variance as long
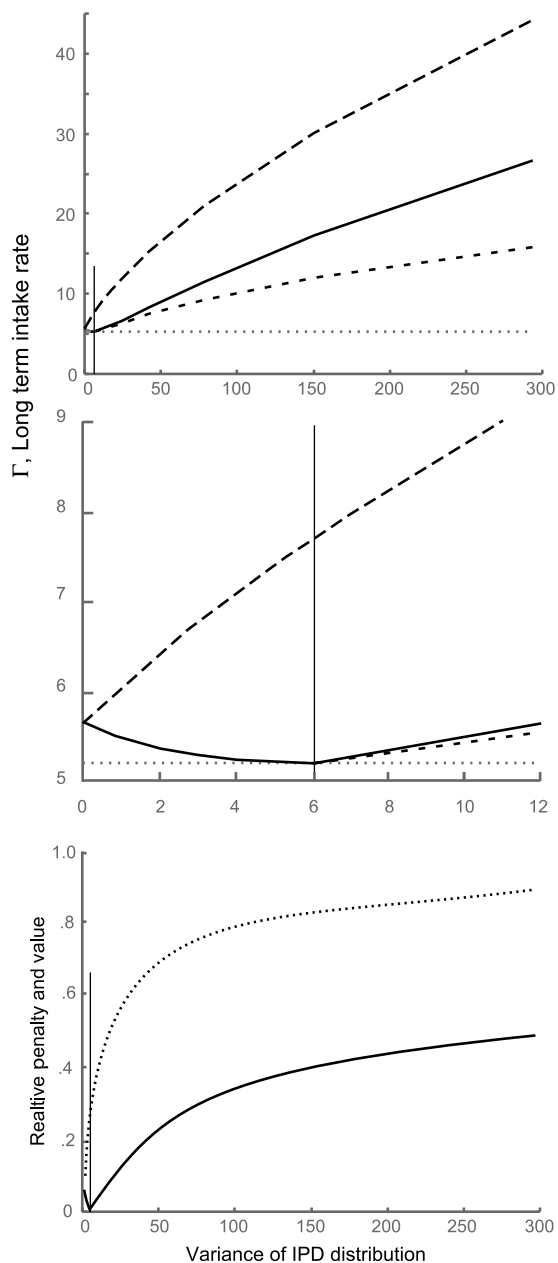
Fig. 5. As in Fig. 4, but variance differs and $\overline{N} = 6$ in all cases. The middle panel is a magnification of the range of low variance-levels. The vertical line indicates the Poisson distribution, as $\overline{N} = s^2$.

as it is less than the mean. However, this results from the skew of the binomial distribution. When the variance is very low the parameter p of the binomial distribution is close to 1. That means that most patches are close to the mean value, and few are poor, this fact benefits the Bayesian foragers capable of estimating instantaneous intake rate. As variance increases the distribution approaches the Poisson. In a Poisson the occurrence of one item gives no information whatsoever about the

occurrence of other items. Therefore, in this case the optimal strategy for a Bayesian forager is to follow a fixed time rule, and the payoff of these two strategies are the same.

Performance of patch use rules under a fixed exogenous cost: scenario 2

In the following we will discuss the second scenario, when the environment has a finite size, a finite time, and the opportunity to save metabolic costs of foraging by resting instead of foraging. The results are quantitatively quite different from the first case (Fig. 6, cf. Fig. 1). The prescient, the Bayesian potential value, and the fixed time foragers all gain their maximum fitness when patches are left when the quitting harvest rate equalizes the foraging costs, $c - c_0$. That is, all three types of foragers leave the patches when they estimate the same patch value, i.e. $\Pi = c - c_0$, despite the fact that their long-term gains are quite different ($\Gamma_{PS} > \Gamma_{BP} > \Gamma_{FT}$ in Fig. 6). In contrast, the instantaneous value Bayesian forager gains its maximum fitness if it uses a quitting harvest rate lower than this, i.e. $\Pi < c - c_0$ (Γ_{BPm} in Fig. 6). If this forager is forced to quit harvesting when $\Pi = c - c_0$ it will gain much less than its maximum (Γ_{BPe} in Fig. 6). This is a direct demonstration that a forager that ignores the FBI when exploiting a clumped distribution is unable to maximize fitness by equalizing the costs and benefits of foraging.

In Fig. 7 we show that the long-term energy gain rate increases, for all three strategies displayed, when the
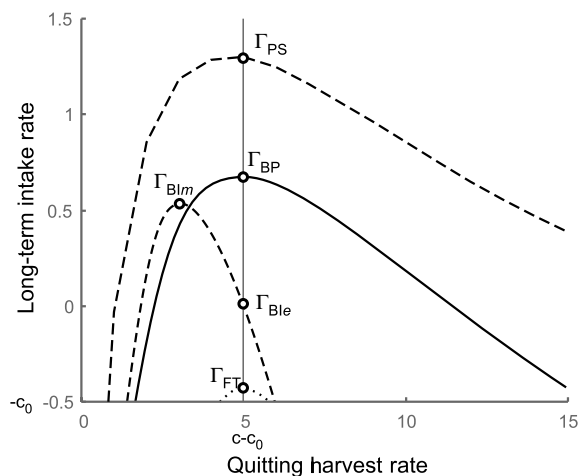

Fig. 6. The fitness for the four different patch use strategies under the fixed exogenous cost model. In all cases $c = 5.5$ and $c_0 = 0.5$ and thus $\Pi^* = c - c_0 = 5$. For the three strategies (PS, BP and FT) the maximum fitness occurs at Π^*. For the Bayesian instantaneous value forager this is not the case. Where $\Pi = c - c_0$ fitness (Γ_{BIe}) is far below the maximum (Γ_{BIm}) for the strategy.
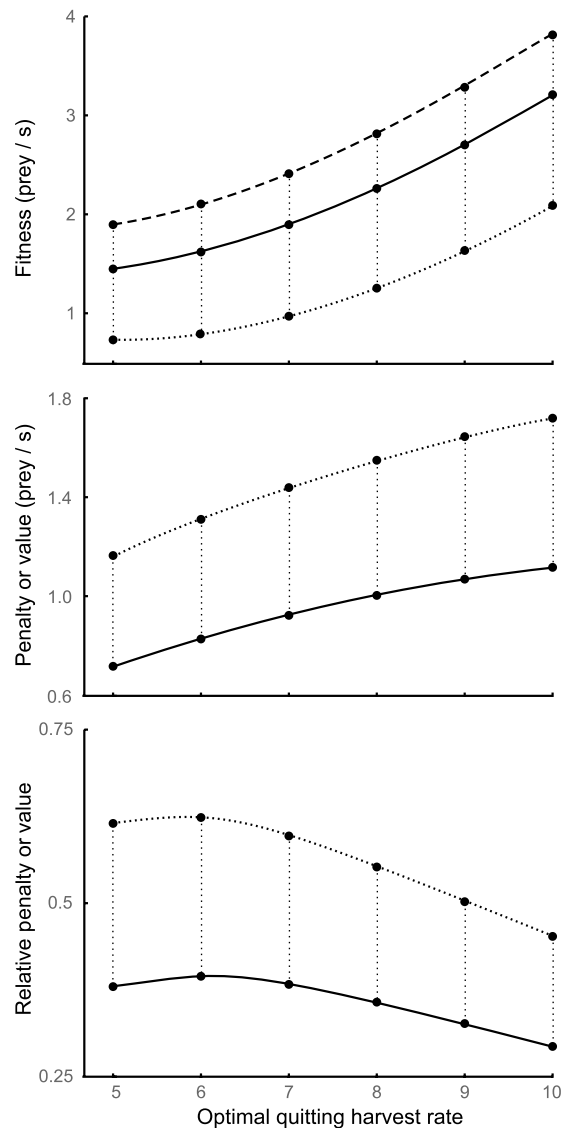
Fig. 7. The top panel shows the fitness when c_0 varies and c is kept constant. $\Pi^* = c - c_0$ varies as a consequence of this. The upper (dashed) curve Γ_{PS}, the middle (solid) curve is Γ_{BP} and the lower (dotted) curve is Γ_{FT}. The middle panel is the penalty of ignorance ($\Gamma_{PS} - \Gamma_{FT}$, dotted curve), and value of sample information ($\Gamma_{BP} - \Gamma_{FT}$, solid curve). The lower panel show these entities relative to Γ_{PS}.

metabolic cost of not foraging decreases, i.e. as $\Pi = c - c_0$ increases. This is straightforward, as the foragers pay less for the time they are inactive. In absolute terms, both the penalty of ignorance and the value of sample information increase with Π, which corresponds with the result that FBI increases with Π (Fig. 3). However, relative to the payoff for the prescient, both the penalty of ignorance and the value of sample information decrease over most of the range. There is a weak increase only for the lowest quitting harvest rate, i.e. for the highest fixed cost of alternative activities.

In Fig. 8 we show the expected giving-up densities in patches over a range of initial prey densities, N. For this figure we use the second scenario, with $c - c_0 = 5$ for all patch use strategies, just as in Fig. 6.

There are two obvious results. First, a prescient forager depletes all patches down to $k = (c - c_0)/A$, if initial prey density $N > k$ (horizontal line PS in Fig. 8). Second, a fixed time forager depletes a constant proportion from all patches, namely $1 - e^{-At^*}$ (FT in Fig. 8).

For the Bayesian strategies, the results are less obvious. In all three cases (Bayesian potential, BP, Bayesian instantaneous equalizing, BI_e, and Bayesian instantaneous maximizing, BI_m) the expected GUD will asymptotically reach their respective Π (i.e. 5, 5 and 3.1 respectively) for high initial prey densities. For lower values of N they will all produce GUDs higher than Π. In the case of BI_e the GUDs will be much higher than Π, and GUD will only be close to Π in extremely rich patches. When H is high relative to $\overline{N}$, the Bayesian forager that ignores FBI leaves patches soon after arrival if prey are not found quickly. This guarantees that the forager only spends lots of time in very rich patches. However, it does so at the cost of discarding most rich and moderately rich patches together with the poor, as can be inferred from Fig. 8.

The Bayesian potential forager, that aims for the same Π, but which recognizes a positive FBI, and hence has $\Pi > H$, stays longer in all patches even without reward (cf. Fig. 2 and 3). By doing so, it wastes a little time in all poor patches, but the forager discards many fewer rich patches, and this outweighs the cost of wasting time in poor patches.

A Bayesian instantaneous forager that maximizes its fitness, and is not forced to leave the patches when $\Pi = c - c_0$, will exploit the patches almost in the manner that the Bayesian potential value forager does, and will also have a fitness not much lower than this strategy. The way it achieves this, however, is by lowering H for all patches. It gains by discarding fewer of the rich patches, but it loses by depleting rich patches unnecessarily low.

The concept of overuse and underuse can apply to these results if we use Π/A as the template GUD that all patches should have (Valone and Brown 1989). An overuse is when the patch has been depleted too far, by spending too much time in it, and underuse is the reverse. It is then easy to see that the prescient is the only strategy succeeding completely, and that the equalizing Bayesian instantaneous value forager underuses practically all patches. The fixed time forager overuses poor patches and underuses all other. The Bayesian potential value forager may be said to overuse poor patches, underuse moderately rich and more or less correctly use very rich patches (Olsson and Holmgren 1998, 2000).

However, it is important to note that the Bayesian foragers are not basing their patch departure on an estimate of prey density per se, but on estimated future prey capture rates. That is, the overuse and underuse by Bayesian foragers does not imply that they commit mistakes in relation to their own strategy – they leave the patches when their Π falls to the threshold level, as stipulated by patch use theory. That is, when the GUDs are viewed as a function of initial prey densities (Fig. 8, Valone and Brown 1989) one may get the impression that Bayesian foragers on average commit mistakes when patches are left. However, when GUDs are viewed as a function of search time (Olsson and Holmgren 1999), which is the state-space relevant to the behavior of the forager, it is clear that on average mistakes are not being made.
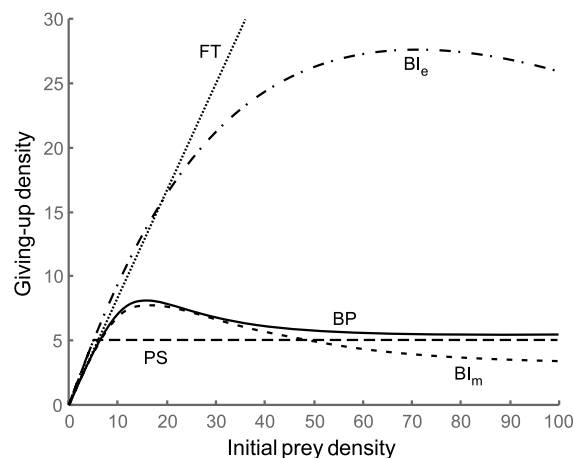

Fig. 8. Giving-up densities (GUDs) as a function of initial prey densities for the four different patch use strategies shown in Fig. 6. The curves for PS, BP, FT, and BIe all use $\Pi = c - c_0$. The curve BIm use $\Pi = 3.1$, which yields the highest fitness for this strategy.

Discussion

In this paper we have shown how a positive foraging benefit of information, FBI, is available to Bayesian foragers exploiting clumped prey. By accepting this benefit the optimal Bayesian foragers (Green 1988, Olsson and Holmgren 1998) reduce their penalty of ignorance, and hence gain a fitness advantage. We have suggested that the general patch use model should be $\Pi = C + P + MOC$ (Brown 1988), or in the case of a rate maximizing forager $\Pi = MOC$ (Charnov 1976). Here $\Pi = H + FBI$, i.e. it is the potential patch quality, including the instantaneous value, H, and the future promises, FBI. It is only when prey distribution is clumped that $FBI > 0$. That is, foragers exploiting prey with Poisson or less aggregated distributions should base their patch departure on $H = C + P + MOC$ (Iwasa et al. 1981). In the case of prescient foragers, patch departure should be based on known number of prey remaining,

which at the same time is a measure of instantaneous value and future promises. Hence, also in that case the original model, $H = C + P + MOC$ is correct (Brown 1988).

Ignoring the FBI when exploiting clumped prey (Iwasa et al. 1981, Valone and Brown 1989, Rodríguez-Gironés and Vásquez 1997) leads to the remarkable result that there is no truly optimal solution for Bayesian foragers, as $H < C + P + MOC$.

This means that, only in one of the cases we have investigated does the correct inclusion of FBI matter. However, we argue that it is a very important case. Very many organisms have a contagious spatial distribution, often because their environment itself is heterogeneous. Among areas where they occur their numbers are often well described by a negative binomial distribution (Pielou 1977, Olsson et al. 1999, van Gils et al. 2003, Vos and Hemerik 2003, Warren et al. 2003). That is, not only is it common to find that organisms have a patchy distribution, occurring in some areas and not in others, but in most of the areas where they do occur they are rare, and only in some are they very common (Warren et al. 2003). Thus, foragers exploiting prey with a negative binomial (or similar) distribution are probably very common in nature.

Most importantly, the foraging benefit of information bridges the gap between the correctness of Green's potential value rule and the incorrect, but appealing cost benefit analyses of patch use. The possibility for good news generates a benefit to sticking around over and above instantaneous expectations of rewards. This reward, FBI, can be measured and placed within the same foraging currencies as any other cost or benefit.

The value of sampling information (multiplied by the marginal rate of substitution of energy for fitness) is the selection gradient acting on completely ignorant foragers, to adopt a Bayesian strategy. Likewise, the penalty of ignorance is the selection gradient for the ignorant to become prescient, and the penalty of sampling information is the selection acting on the Bayesian to become prescient. The step from ignorant to Bayesian may involve mainly neural capacity, as most animals probably already have some estimates of both time and food intake/gut fullness. However, the neural capacity that is needed by a Bayesian forager is not great (Holmgren and Olsson 2000), the question may rather be if there is a cost to it. The evolution of prescience may involve the development of new sensory abilities, such as olfaction. In many cases it may not be that the senses are lacking, but rather that the information it conveys is ambiguous, and hence not valuable.

It is easy to see that these selection pressures will be very different in different environments (Fig. 4 and 5). The penalty of ignorance is relatively insensitive to variations in mean prey density. However, the selection to evolve from ignorant to Bayesian is stronger the richer the environment, and the selection to evolve to prescient is weaker. The variance of the prey distribution dramatically increases the selection for an ignorant forager to evolve prescience, particularly at low and intermediate variances. Interestingly, the selection for evolving Bayesian capacity decreases with variance for dispersed distributions, and is entirely absent for Poisson, but then increases strongly for clumped distributions.

The above predicts that we should mostly find Bayesian foragers among animals that prey on heavily clumped food. As far as we know, the natural prey distribution has only been estimated in two of the cases when a Bayesian strategy has been found (Olsson et al. 1999, van Gils et al. 2003). In both of these cases the distribution was clearly clumped, with the over dispersion coefficient close to 0.3, which conforms to the prediction.

Empirical tests of patch use strategies

Bayesian foraging strategies have been reported in a number of cases in natural or experimental settings (Valone and Brown 1989, Valone 1991, Alonso et al. 1995, Olsson et al. 1999, Davidson and Morris 2001, van Gils et al. 2003, reviewed by Valone 2006).

Whenever dealing with clumped prey distributions, we would like to advocate the use of tests of the positive relation between GUDs (or instantaneous quitting harvest rates) and search time (Olsson et al. 1999, van Gils et al. 2003). This correlation is only predicted for the Bayesian potential value forager (Fig. 2 and 3), and it is based on the same dimensions (harvest rate by search time) as is the foragers' own patch departure decision.

Other predictions that have been tested are the relation between GUDs and initial prey densities, and the expectation that initially rich patches should be underused and poor patches overused. These predictions are in a sense secondary properties, as they do not directly measure the variables that the foragers may have based their decisions on. Also, they are more ambiguous as all strategies are predicted to have an initially increasing relation between GUD and initial prey density. And both the Bayesian potential strategy and the prescient should level off, and they may be difficult to separate. In addition, this relation for the Bayesian is non-linear, and cannot be described with a simple analytic equation (Olsson and Holmgren 2000). A way to partly overcome this is to use polynomial regression to test for the expected non-linearities (Davidson and Morris 2001). Still, the range of prey densities that may need to be used in an experiment will make it difficult to accomplish.

An additional point, worth making here, is that the Bayesian potential and Bayesian instantaneous rate rules

need never be separated in a test. The Bayesian instantaneous rate rule should simply not be expected, as it is not only sub-optimal – it is incorrect.

Perspectives beyond foraging

So far, explicit population or community level effects have not been modeled for Bayesian foragers. Some initial predictions on this theme have been presented by Rodríguez-Gironés and Vásquez (1997) and Olsson and Holmgren (2000). Their conclusions were that when foragers use a Bayesian strategy, the distribution of prey among patches will have profound effects on the stability of the system. However, it was also predicted that the currency of the predator might influence density dependent mortality rate of the prey. The conclusion to be made now is that the Bayesian foraging models have reached the point where they are sufficiently theoretically and empirically founded, and practically manageable, to be included in population and community models. Progress in this direction would certainly be worthwhile.

Bayesian decision-making is likely important in many behaviors other than foraging. For example, the assessment a male is making of its rivals (Peake et al. 2002) or a female is making of potential mates (Luttbeg 1996, Mennill et al. 2002) may be Bayesian. This study hints that such assessments and decisions will be particularly intriguing when the two contestants have different reward distributions. For example, it seems plausible, that females could have a right-skewed distribution and males a left-skewed, in a mate choice situation.

References

Alonso, J. C., Alonso, J. A., Bautista, L. M. et al. 1995. Patch use in cranes: a field test of optimal foraging predictions. – Anim. Behav. 49: 1367–1379.

Brown, J. S. 1988. Patch use as an indicator of habitat preference, predation risk, and competition. – Behav. Ecol. Sociobiol. 22: 37–47.

Brown, J. S. 1992. Patch use under predation risk. I. Models and predictions. – Ann. Zool. Fenn. 29: 301–309.

Brown, J. S. and Mitchell, W. A. 1989. Diet selection on depletable resources. – Oikos 54: 33–43.

Caraco, T. 1980. On foraging time allocation in a stochastic environment. – Ecology 61: 119–128.

Charnov, E. L. 1976. Optimal foraging, the marginal value theorem. – Theor. Popul. Biol. 9: 129–136.

Davidson, D. L. and Morris, D. W. 2001. Density-dependent foraging effort of deer mice (*Peromyscus maniculatus*). – Funct. Ecol. 15: 575–583.

Fretwell, S. D. and Lucas, H. L. J. 1970. On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. – Acta Biother. 19: 16–36.

Green, R. F. 1980. Bayesian birds: a simple example of Oaten's stochastic model of optimal foraging. – Theor. Popul. Biol. 18: 244–256.

Green, R. F. 1984. Stopping rules for optimal foragers. – Am. Nat. 123: 30–40.

Green, R. F. 1988. Optimal foraging for patchily distributed prey: random search.-Tech. Rep. 88-2. Dept of Mathematics & Statistics, Univ. of Minnesota, Duluth.

Gross, J. E., Zank, C., Hobbs, N. T. et al. 1995. Movement rules for herbivores in spatially heterogeneous environments-responses to small-scale pattern. – Landscape Ecol. 10: 209–217.

Holmgren, N. M. A. and Olsson, O. 2000. A three-neuron model of information processing during Bayesian foraging.-In: Malmgren, H., Borga, M. and Niklasson, L. (eds), Perspectives in neural computing: artificial neural networks in medicine and biology. Proc. ANIMAB-1 Conference, Göteborg, Sweden, May, 2000. Springer, pp. 265–270.

Iwasa, Y., Higashi, M. and Yamamura, N. 1981. Prey distribution as a factor determining the choice of optimal foraging strategy. – Am. Nat. 117: 710–723.

Luttbeg, B. 1996. A comparative Bayes tactic for mate search assessment and choice. – Behav. Ecol. 7: 451–460.

McNamara, J. M. 1982. Optimal patch use in a stochastic environment. – Theor. Popul. Biol. 21: 269–288.

Mennill, D. J., Ratcliffe, L. M. and Boag, P. T. 2002. Female eavesdropping on male song contests in songbirds. – Science 296: 873.

Mitchell, W. A. 1989. Informational constraints on optimally foraging hummingbirds. – Oikos 55: 145–154.

Oaten, A. 1977. Optimal foraging in patches: a case for stochsticity. – Theor. Popul. Biol. 12: 263–285.

Olsson, O. and Holmgren, N. M. A. 1998. The survival-rate-maximizing policy for Bayesian foragers: wait for good news. – Behav. Ecol. 9: 345–353.

Olsson, O. and Holmgren, N. M. A. 1999. Gaining ecological information about Bayesian foragers through their behaviour. I. Models with predictions. – Oikos 87: 251–263.

Olsson, O. and Holmgren, N. M. A. 2000. Optimal Bayesian foraging policies and prey population dynamics-some comments on Rodríguez-Gironés and Vásquez. – Theor. Popul. Biol. 57: 369–375.

Olsson, O., Wiktander, U., Holmgren, N. M. A. et al. 1999. Gaining ecological information about Bayesian foragers through their behaviour. II. A field test with woodpeckers. – Oikos 87: 264–276.

Peake, T. M., Terry, A. M. R., McGregor, P. K. et al. 2002. Do great tits assess rivals by combining direct experience with information gathered by eavesdropping? – Proc. R. Soc. B 269: 1925–1929.

Pielou, E. C. 1977. Mathematical ecology. – Wiley-Interscince.

Real, L. A. 1980. Fitness, uncertainty, and the role of diversification in evolution and behavior. – Am. Nat. 115: 623–638.

Rodríguez-Gironés, M. A. and Vásquez, R. A. 1997. Density-dependent patch exploitation and acquisition of environmental information. – Theor. Popul. Biol. 52: 32–42.

Rosenzweig, M. L. 1981. A theory of habitat selection. – Ecology 62: 327–335.

Sallabanks, R. 1993. Hierarchical mechanisms of fruit selection by an avian frugivore. – Ecology 74: 1326–1336.

Sandlin, E. A. 2000. Cue use affects resource subdivision among three coexisting hummingbird species. – Behav. Ecol. 11: 550–559.

Stephens, D. W. 1989. Variance and the value of information. – Am. Nat. 134: 128–140.

Valone, T. J. 1991. Bayesian and prescient assessment: foraging with pre-harvest information. – Anim. Behav. 41: 569–577.

Valone, T. J. 2006. Are animals capable of Bayesian updating? An empirical review. – Oikos 112: 252–259.

Valone, T. J. and Brown, J. S. 1989. Measuring patch assessment abilities of desert granivores. – Ecology 70: 1800–1810.

Valone, T. J. and Giraldeau, L.-A. 1993. Patch estimation by group foragers: what information is used. – Anim. Behav. 45: 721–728.

van Gils, J. A., Schenk, I. W., Bos, O. et al. 2003. Incompletely informed shorebirds that face a digestive constraint maximize net energy gain when exploiting patches. – Am. Nat. 161: 777–793.

Vos, M. and Hemerik, L. 2003. Linking foraging behavior to lifetime reproductive success for an insect parasitoid: adaptation to host distributions. – Behav. Ecol. 14: 236–245.

Warren, M., McGeoch, M. A. and Chown, S. L. 2003. Predicting abundance from occupancy: a test for an aggregated insect assemblage. – J. Anim. Ecol. 72: 468–477.

Subject Editor: Per Lundberg