LUND UNIVERSITY

Bayesian foraging with only two patch types

Olsson, Ola

Link to publication

Bayesian foraging with only two patch types

Ola Olsson

Olsson, O. 2006. Bayesian foraging with only two patch types. – Oikos 112: 285–297.

I model the optimal Bayesian foraging strategy in environments with only two patch qualities. That is, all patches either belong to one rich type, or to one poor type. This has been a situation created in several foraging experiments. In contrast, previous theories of Bayesian foraging have dealt with prey distributions where patches may belong to one out of a large range of qualities (binomial, Poisson and negative binomial distributions). This study shows that two-patch systems have some unique properties. One qualitative difference is that in many cases it will be possible for a Bayesian forager to gain perfect information about patch quality. As soon as it has found more than the number of prey items that should be available in a poor patch, it "knows" that it is in a rich patch. The model generates at least three testable predictions. 1) The distribution of giving-up densities, GUDs, should be bimodal in rich patches, when rich patches are rare in the environment. This is because the optimal strategy is then devoted to using the poor patches correctly, at the expense of missing a large fraction of the few rich patches available. 2) There should be a negative relation between GUD and search time in poor patches, when rich patches are much more valuable than poor. This is because the forager gets good news about potential patch quality from finding some food. It therefore accepts a lower instantaneous intake rate, making it more resistant against runs of bad luck, decreasing the risk of discarding rich patches. 3) When the energy gains required to remain in the patch are high (such as under high predation risk), the overuse of poor patches and the underuse of rich increases. This is because less information about patch quality is gained if leaving at high intake rates (after short times). The predictions given by this model may provide a much needed and effective conceptual framework for testing (both in the lab and the field) whether animals are using Bayesian assessment.

O. Olsson, Dept of Animal Ecology, Lund Univ., Ecology Building, SE-223 62 Lund, Sweden (ola.olsson@zooekol.lu.se).

Theoretical treatments of Bayesian foraging have mostly focussed on prey distributions where the number of prey present per patch may take any of a large (or infinite) number of values (binomial, Poisson and negative binomial distributions, here called smooth distributions; Green 1980, 1984, 1987, 1988, Iwasa et al. 1981, Olsson and Holmgren 1998, 2000, Olsson and Brown 2006), although some studies have been general (Oaten 1977, McNamara 1982, Green 2006).

Experimental studies have often focussed on a contrast between a rich patch and a poor patch (Lima 1984, Valone and Brown 1989, Valone 1992, Meyer and Valone 1999, Vásquez et al. 2006). That is,

they have created a system with only two patch types (or sub-types sensu Stephens and Krebs 1986). This is what I here will call a two-patch system. Many general aspects of Bayesian foraging behaviour that are known from theoretical treatments of smooth distributions are probably applicable to a two-patch system. However, a two-patch system also has some properties that are qualitatively different from smooth distributions. For example, if there is one patch type that always has 10 prey items, and another with 20 items, the forager has perfect information about patch quality once it has taken 11 items or more. A forager exploiting e.g. a Poisson or negative binomial

distribution will never be able to gain perfect information about patch quality.

At least since Iwasa's et al. (1981) seminal paper, it is well known that different smooth distributions give the estimate of instantaneous intake rate (r) different properties (Olsson and Brown 2006). In a binomial distribution r declines with the number of prey taken, and with search time. In a Poisson distribution r declines with search time but is independent of prey found. In a negative binomial distribution r declines with search time and increases with the number of prey found. Thus, in smooth distributions, the behaviour of r depends on whether the variance in prey density is less than, equal to, or greater than the mean of the prey density distribution. It is easy to generate two-patch distributions with the variance less than, equal to, or greater than the mean, but the properties of r in such distributions are not known.

In this paper I will present a model of the optimal Bayesian patch leaving rules when exploiting two-patch systems. The model generates some important, and even surprising, predictions that are quite different from those generated by models for unimodal distributions. Two-patch systems are sometimes found in nature when the qualities of patches may be in two more or less distinct states. However, the greatest value of this model is probably that it suggests promising and convincing ways for experimentally testing whether animals are capable of using Bayesian strategies, or not.

Model

The model is an application of Green's (2006) general model. Here, I have made the extensions necessary to make it apply to random search and a two-patch system.

I assume that the forager "knows" the prey density distribution (the prior) in the environment and is capable of Bayesian estimation of the patches it exploits. The forager then only needs to keep track of search time spent (t) in a patch and number of prey found (n) to make the estimations necessary.

The "world" contains only two types of patches. A poor (type 1) and a rich (type 2). I will denote the initial prey density by N and use the index j to identify patch types. Thus, initial prey densities are N_1 and N_2 and $N_1 < N_2$.

The probability to encounter a patch of type j is the prior distribution, Pr_j. As there are only two patch types $Pr_2 = 1 - Pr_1$.

The posterior distribution is the probability that the patch is of type 1 or 2, conditional on some sampling information. It is found by applying Bayes theorem. With this distribution, the posterior probability that the patch is of type j, when n items have been found after having spent t time steps searching, is then:

$$Po_j = \frac{x_j}{x_1 + x_2} \tag{1}$$

where

$$x_j = Pr_j \binom{N_j}{n} q^{N_j - n} \qquad \text{for } 0 \leq n \leq N_j$$
$$x_j = 0 \qquad \qquad \text{for } n > N_j \tag{2}$$

and

$$q = e^{-At} \tag{3}$$

where A is the searching efficiency.

The instantaneous intake rate (or equivalently the current prey density, if A = 1) is then

$$r_{n,t} = A(Po_1(N_1 - n) + Po_2(N_2 - n)) \tag{4}$$

Equation 2 and 4 show, what was stated above, that the forager's estimate of patch quality gets perfect once n > N_1, with this distribution.

As pointed out elsewhere (Green 1980, 1984, 1987, 1988, 2006, Olsson and Holmgren 1998, 1999, 2000, Olsson and Brown 2006), a Bayesian forager should not base its decision to leave the patch on the estimate of instantaneous intake rate, or current prey density. Instead, a patch should be left when the potential intake rate (Π) during the remainder of the patch visit equals some target rate, C. This rate could be the long-term intake rate (Green 2006), but that is not necessary (Olsson and Holmgren 1998, Olsson and Brown 2006).

The potential intake rate is the expectation of prey capture during the rest of the stay in the patch, given that the forager follows the optimal departure rule. It is hence expressed as the ratio of the expected number of prey to find, to the expected time to spend before the patch is left (Green 1980, 1984, 1987, 1988, 2006, Olsson and Holmgren 1998, 2000, Olsson and Brown 2006). Both the expected number of prey, G, and the expected search time, S, are found by dynamic programming, by assuming that the patch must be left by some final time T. In order to calculate Π, we need some more numbers, which are given by Eq. 5 through 8.

The probability of finding i items in the interval Δ, if the patch has $N_j - n$ items when the interval starts, is:

$$P(i, N_j - n) = \binom{N_j - n}{i} m^i k^{N_j - n - i} \tag{5}$$

(Olsson and Holmgren 2000, Green 2006) where k = $e^{-A\Delta}$ and m = 1 - k.

Finally, ρ is calculated as

$$\rho(i, t, n) = \sum_{j=1}^{2} Po_j P(i, N_j - n) \tag{6}$$

Green (2006) suggests that

$$G(t, n) = \sum_{i=0}^{N_2 - n} \rho(i, t, n)(i + G(t + \Delta, n + i)) \qquad (7)$$

and

$$S(t, n) = \Delta + \sum_{i=0}^{N_2 - n} \rho(i, t, n)(S(t + \Delta, n + i)) \qquad (8)$$

where i are the number of prey items to be found in the next time step, Δ. In Eq. 7 and 8, N_{max} (Green 2006) is replaced by N_2, as that is the maximum in this case.

The probability of finding i items in the next time step, $\rho(i,t,n)$, is conditional on that t time steps have been spent searching, and n prey items have been found.

Then, the ratio $\Pi = G/S$ is evaluated for each n and t, and the forager should remain in the patch as long as $\Pi > C$ (McNamara 1982, Green 2006).

Results

I evaluated the model for several different two-patch distributions, and critical levels of potential intake rate, C. The purpose of the analysis is both to analyze Bayesian foraging in two patch systems, which has not been done before, and to generate testable predictions. Testable predictions for a two-patch system may be among the strongest tools to identify Bayesian foraging in experimental settings. This is both because such systems are elegantly simple, and because the predictions generated are often quite distinct from those generated by alternative models.

In a two-patch system mean prey density per patch is $\tilde{N} = Pr_1 N_1 + Pr_2 N_2$. The variance in prey density is thus $s^2 = Pr_1(\tilde{N} - N_1)^2 + Pr_2(\tilde{N} - N_2)^2$.

There are two principal ways (set A and set B; Table 1) of creating two-patch distributions with different var-iance-to-mean ratios. Firstly, I create a set of distributions (set A) by keeping the values of the patches constant ($N_1 = 10$, $N_2 = 20$), and vary Pr_j, so that the variance is less than, equal to or larger than the mean.

In the second set (B), I fix the mean at $\tilde{N} = 15$, and $Pr_1 = Pr_2 = .5$. Then I let N_1 and N_2 vary such that I get distributions with the variance less than, (almost) equal to, and larger than the mean (here $N_2 = 30 - N_1$).

The baseline case is the one joining the two sets, with $N_1 = 10$, $N_2 = 20$, and $Pr_1 = Pr_2 = .5$. All parameter combinations are shown in Table 1.

Instantaneous intake rates

In Fig. 1 I show how the estimated instantaneous intake rate (or current prey density, as A = 1) depends on search time and number of prey items found, in the different two-patch distributions. There are two common features for all these distributions that are worth noting. They distinguish two-patch distributions from other distributions. Firstly, as mentioned above, once n is higher than N1 + 1, the estimate becomes independent of search time, and only depends on the number of prey found.

Focus on the left column of Fig. 1 (simulations 1-4). When n = 11, r = 9, when n = 12, r = 8 and so on. That is, the forager has perfect information about patch quality in this region. In the region where zero to ten items are caught, and long search times have been spent, the estimate is seemingly close to being perfect, too. If e.g. 10 items have been found after two time units, r is very close to zero. However, the forager is fully prepared to change its mind, should it find an 11th item. Then r jumps up to nine (the likelihood of doing so, with that sampling experience is however almost nil). This is the second feature.

Table 1. The parameter values used in the 11 simulations. N is the number of prey initially in the patches, Pr is the frequency of poor and rich patches respectively, and C is the critical potential intake rate when the forager should leave the patch. Set A all have the same types of patches, but in varying frequencies (hence varying mean and variance). Set B all have the same mean, but varying patch types (hence varying variance). Simulation 1 is a baseline case, which is common for both sets. Simulations 5 and 6 are the baseline case, but with different C-values. In all simulations T = 2.5 and Δ = 0.01. A = 1.

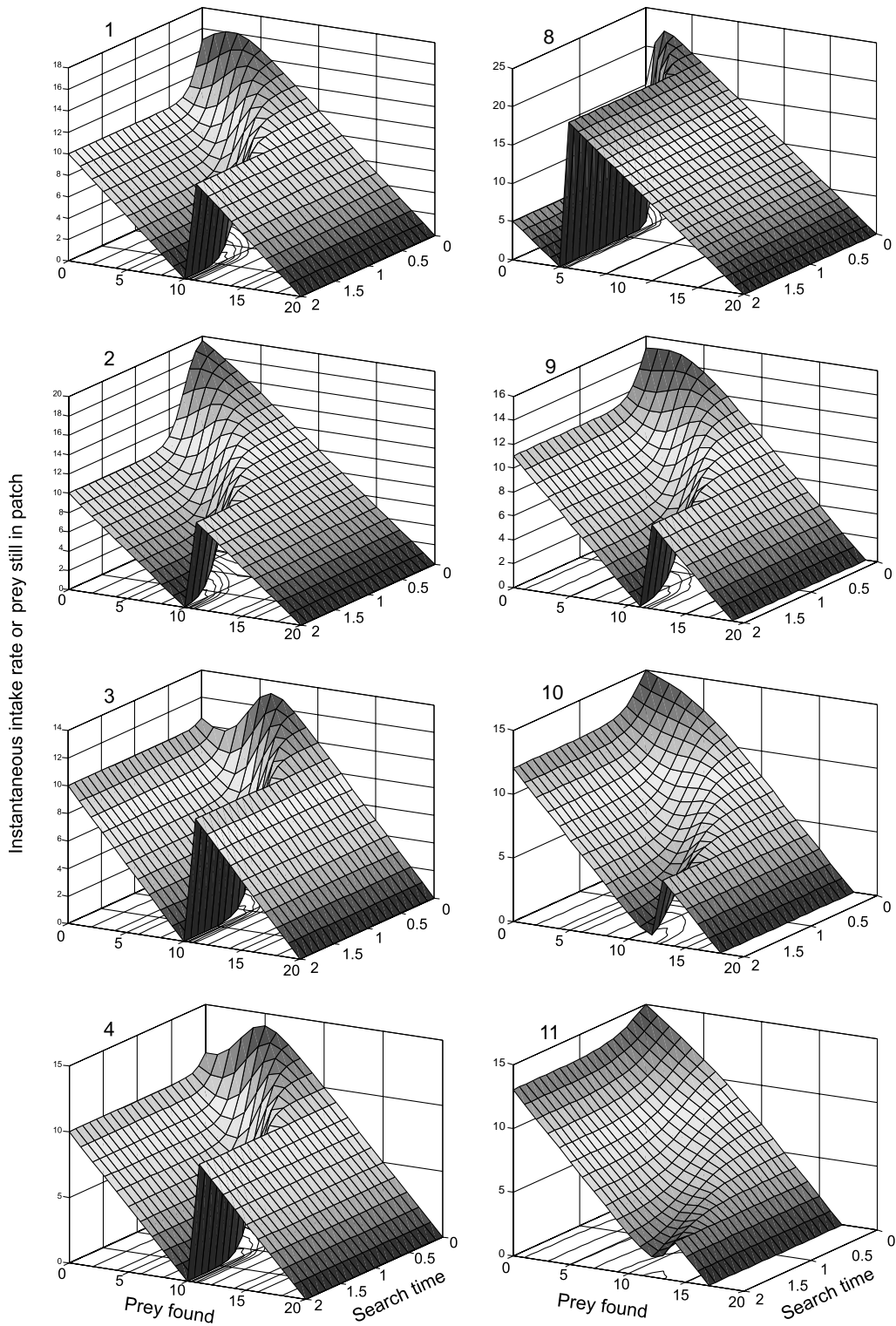|		Parameters						
		N		Pr				
Set	Simulation	poor	rich	poor	rich	C	Mean	Variance
A, B	1	10	20	0.50	0.50	7	15	25
A	2	10	20	0.05	0.95	7	19.5	4.75
A	3	10	20	0.95	0.05	7	10.5	4.75
A	4	10	20	0.87	0.13	7	11.3	11.3
A, B	5	10	20	0.50	0.50	2	15	25
A, B	6	10	20	0.50	0.50	12	15	25
B	7	1	29	0.50	0.50	7	15	196
B	8	5	25	0.50	0.50	7	15	100
B	9	11	19	0.50	0.50	7	15	16
B	10	12	18	0.50	0.50	7	15	9.0
B	11	13	17	0.50	0.50	7	15	4.0

Fig. 1. The estimated instantaneous intake rate ($r_{n,t}$) as a function of number of prey caught (n) and search time spent in the patch (t). The forager arrives in the patch with n = 0 and t = 0, i.e. in the inner corner of the graphs. As it spends time and finds food, the estimate of patch quality changes. The numbers in each panel refers to the number of the simulation. See Table 1 for parameter values.
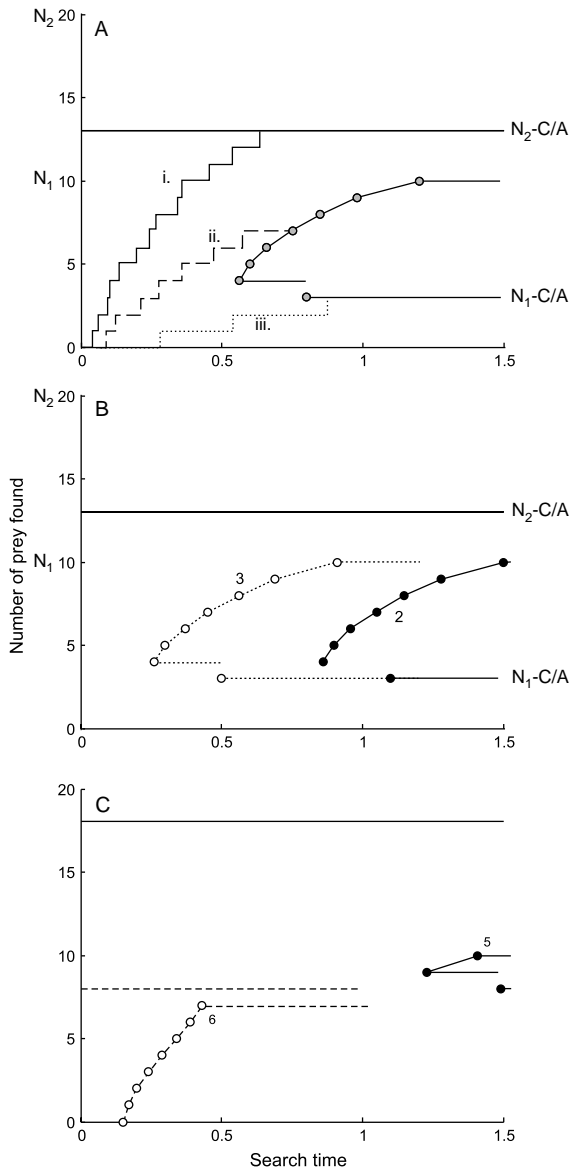
Fig. 2. Stopping points for Bayesian foragers exploiting two-patch distributions with parameters given in Table 1. Panel (A) shows the stopping points for simulation 1 with three specific examples (i, ii, and iii) drawn. When a forager arrives in a new patch its search time is zero, and it has found zero prey. Thus it appears in the origin. As it searches it moves to the right in the graph, and for every prey item it finds it moves one step upwards. Its estimate of patch quality changes accordingly (Fig. 1). The forager shown by line i is quite successful and has found $N_2 - C/A = 13$ items after approximately 0.6 time units. It then knows that the patch was initially rich (N_2) and now has $C/A = 7$ items left. Hence, it has reached its threshold and the patch should be left, as indicated by the solid horizontal line at $N_2 - C/A$. The forager shown by the dotted line iii is much less successful. It finds only three items in about 0.9 time units. With this experience it is quite certain that the patch is of quality N_1 and only $C/A = 7$ items are left. Hence, it is time to leave the patch, as it has hit the horizontal line at $N_1 - C/A$. The forager shown by the dashed line ii has intermediate success. It finds 7 items in about 0.7 time units. Here the forager cannot be very certain about the quality of the patch, it may be rich or poor,

Mostly, the estimate of patch quality declines for every item found (the forager thinks there is less left, for each item that is taken), but there are regions where r increases in n (the forager thinks it has hit a rich patch). For other distributions (binomial, Poisson, negative binomial) r is always a monotonic function of n (Iwasa et al. 1981). Two-patch systems are therefore qualitatively different to exploit from any other distribution, as the estimate of patch quality may go either up or down when more prey is found (Fig. 1).

As the estimate of patch quality may increase with finding prey, the foraging benefit of information (FBI: Olsson and Brown 2006) is positive when exploiting two-patch systems. Then information is itself worth foraging for, as the information might tell the forager that the patch is richer than it presently estimates. Hence, it is crucial to base patch departure on the potential, rather than instantaneous intake rate in such distributions (Green 1988, Olsson and Holmgren 2000, Olsson and Brown 2006). Basing the departure rule on instantaneous rate (which is shown in Fig. 1) misses the promises for the future that are contained in the sampling information gained.

There are interesting differences between the eight distributions shown in Fig. 1. The distributions are generated such that they have variances less than, equal to, or greater than the mean values. These are properties that they share with the binomial, Poisson, and negative binomial distributions, respectively. In a binomial distribution, r decreases with n. This is also true in a two-patch distribution with the variance less than the mean (mimicking a binomial distribution), when the forager has just arrived in the patch. Similarly, when variance and mean are equal, r is initially independent of n, just as in a Poisson distribution. When variance is greater than the mean, r increases with n, as in a negative binomial distribution. However, these similarities with the smooth distributions hold only when the forager has just arrived in the patch. After some more time spent, and/or prey caught the r shifts into regions where it may either increase or decrease with n.

Stopping points

The patch departure rule can be described by a set of [n, t]-stopping points (Fig. 2, 3), and one or more lines with

but with the experience of the patch the odds are that it is poor and should be abandoned. To the left of the stopping point the potential intake rate is still above C, but to the right of the point it is below. The strategy is to leave the patch at the stopping point. Common for all the stopping points and the horizontal lines is that here the potential intake rate, $\Pi = C$. In (A) and (B) simulations 1-3 are shown. They differ with respect to environmental abundance of rich and poor patches – simulation 2 being the richest environment. In (C) simulations 5 and 6 are shown. They are both the same distribution as simulation 1, but with different critical potential intake rates (C), that governs patch departure.
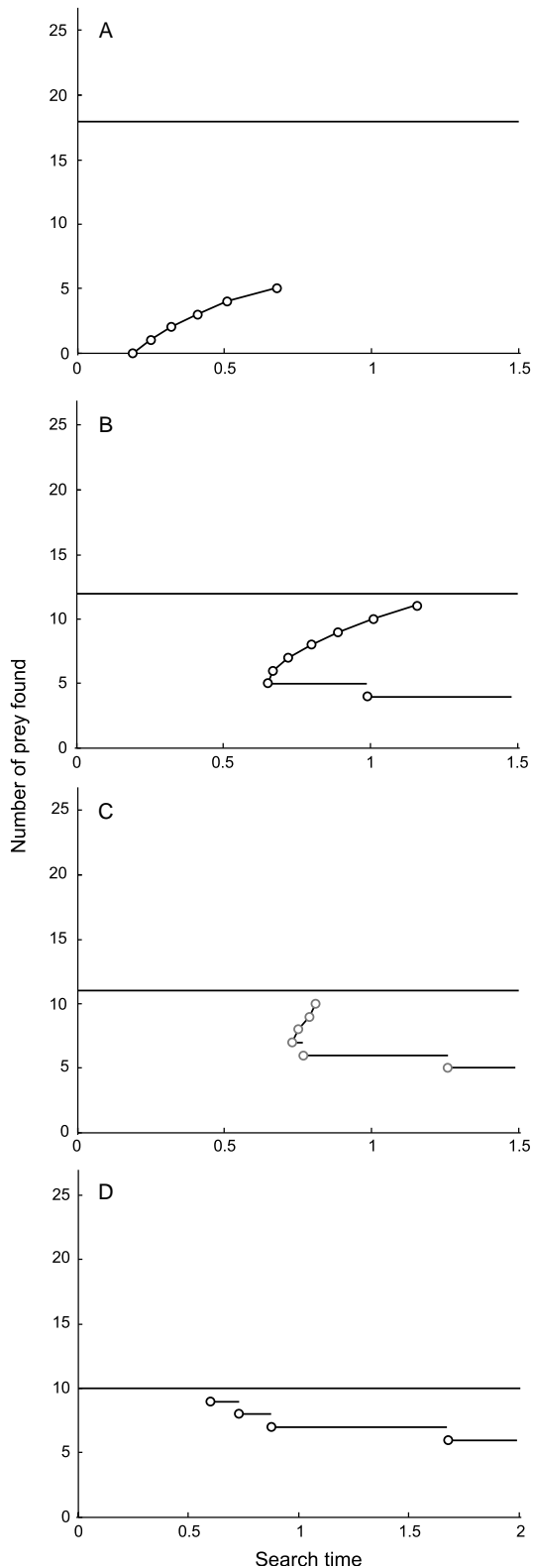
Fig. 3. As in Fig. 2 for simulations 8 (A), 9 (B), 10 (C), and 11 (D).

constant n. At each of these stopping points the potential intake rate equals the same critical value, i.e. $G/S = C$.

As long the forager is below the lines or to the left of the points, it should remain in the patch. Figure 2A shows the baseline case with three examples, which are described in more detail in the figure legend. The uppermost line is where the forager has found $N_2 - C/A$ prey items. It has perfect information, and "knows" that it is using a rich patch, and should leave when the patch is depleted to a constant instantaneous quitting harvest rate (Charnov 1976, Brown 1988). Forager i in Fig. 2A leaves the patch because it has hit this line from below.

At the lower line, where the forager has found $N_1 - C/A$, the patch is almost (but not quite) certainly poor. Forager iii leaves the patch because it hits this line from below. This region of the stopping rule is similar to that for a binomial distribution (Green 1988).

When the forager has found a number of prey items that is $N_1 - C/A \leq n \leq N_1$ (and has not yet passed to the right of any stopping point) the forager should spend more time in the patch the more prey it has found. This is because finding prey here indicates that the patch may be more valuable than initially thought, i.e. FBI is positive (Fig. 1). Forager ii leaves the patch because it hits one of the stopping points (at $n = 7$) from the left.

The general shape of the stopping rule is similar for the different two-patch distributions in set A. It is shifted to the right for more rewarding distributions (Fig. 2B), and down when the critical intake rate to remain in the patch (C) is high (Fig. 2C).

The stopping points (including the two horizontal lines at $N_1 - C/A$ and $N_2 - C/A$) delineates the region of the [n, t]-space in which the forager should remain in the patch. Of course, the forager can only "move" in the step wise manner described by the examples i, ii and iii as prey items come in integer values, and finding a prey item is assumed to be instantaneous. Also, the lines $N_1 - C/A$ and $N_2 - C/A$ goes on to infinity, so given that the forager has found 0, 1 or 2 prey and is below line $N_1 - C/A$ or has found 11 or 12 items and is below line $N_2 - C/A$ it will eventually find enough prey to cross the line. Before it does so, the patch is too rich to abandon.

When $N_2 - C/A \leq N_1$ (simulations 6, 10, 11) the forager will never reach perfect information, as it will always leave before it has taken $N_1 + 1$ items. For the other simulations, the optimal policy is to take more prey from the rich patch than is available in the poor. Thus, once $N_1 + 1$ items are taken the forager gets perfect information about patch quality. When the forager does not have a chance to reach perfect information it will need to base the use of both patch types entirely on sampling experience. This may lead to overuse of poor patches, and underuse of rich (Fig. 6, 7, Valone and Brown 1989, Meyer and Valone 1999).

Simulations 1-5 and 9 contrasts with simulations 6, 7, and 8 in the relative value of the poor patch type. In the former group the poor type is richer than the threshold level, i.e. $N_1 > C/A$. Thus, every patch is valuable and should, even with perfect information, be exploited for some time. In the latter group poor patches have $N_1 < C/A$ and so should ideally be skipped over. This can be seen from the stopping rules in Fig. 2C (simulation 6) and Fig. 3A. If no prey are found within some time period the patch is abandoned. The ideal for a forager with perfect information would be to abandon these patches without searching them at all, but a Bayesian forager can only identify them by sampling.

For simulation 6, the policy looks quite similar to that for a negative binomial distribution, but with a ceiling at the fixed $n = N_2 - C/A$. The maximum patch quality is fixed at N_2, rather than open ended as in the negative binomial. Also, in a negative binomial distribution, the stopping points form a concave, rather than convex curve. A fact worth noting is that all the different stopping rules have a similar shape in set A, and their shapes do not depend on the variance-to-mean ratio.

However, for set B (when mean is fixed at 15, and both patch types are equally common) the general shape of the stopping rules does seem to depend on the variance to mean ratio (Fig. 3). When the variance is 100 ($N_1 = 5$, $N_2 = 25$), the stopping points have a positive slope. This is also true for a range of the points in the baseline case (Fig. 2A), which is part of set B with variance 25, and when variance is 16 (Fig. 3B). When the variance is only 4 (Fig. 3D) the stopping points instead have a negative slope, as when exploiting a binomial distribution. When the variance is 9 ($N_1 = 12$, $N_2 = 18$; Fig. 3C) the relation between n and t at the stopping points is somehow sigmoid. To generate a two-patch distribution with the variance equal to the mean, with $Pr_1 = Pr_2$ and $\tilde{N} = 15$, one would need $N_1 = 11.12$, and $N_2 = 18.88$. Of course, the prey numbers have to be integers, so this distribution is not possible. The two distributions with variance 9 and 16 are the ones closest to an equal variance-to-mean ratio (Fig. 3B-C). That is, they most closely mimic that property of a Poisson distribution. A Bayesian forager exploiting a Poisson distribution should use a fixed time strategy, but when they exploit a two-patch distribution with $s^2 \approx \tilde{N}$ this is not the case. The reason for this is probably that a forager can always gain information from a two-patch distribution. It cannot from a Poisson distribution (Iwasa et al. 1981, Olsson and Brown 2006).

The break point between when the stopping points form a positive and a negative curve thus does not seem to occur precisely when $s^2 = \tilde{N}$. However, the exact shape of the stopping points probably depends on the value of C, in combination with the parameters of the distribution, and I do not intend to investigate this further in the present study.

Quitting harvest rates

An optimal Bayesian forager should leave the patches when the potential intake rate (G/S) equals some constant critical value (here C). At these points the forager will also estimate some instantaneous intake rate, but this instantaneous rate need not be the same in all points (Olsson and Holmgren 1998). In fact as has been shown by Olsson and Brown (2006) the patches should be left when $C = H + FBI$, i.e. when the sum of the instantaneous harvest rate (H) and the foraging benefit of information equals the potential quitting harvest rate. FBI might be zero or positive, but not negative, and H will thus be equal to or smaller than C. In this model I assume that the forager searches randomly, and hence the number of prey left in a patch (the giving-up density, GUD) is simply H/A.

When there is a large difference between the rich and poor patch, the quitting harvest rates deviate considerably from a constant value. In the most extreme distribution, where the rich patch has five times more prey items than the poor, the quitting harvest rate decreases almost linearly with search time (the solid line in Fig. 4A). These quitting harvest rates correspond to the positively related stopping points, for that distribution (Fig. 3A). It should be noted that most rich patches will be left when $n = N_2 - C/A$, i.e. when $H = C$ (the dashed line in Fig. 4A). In fact, for this distribution, only about 10^{-5} of the rich patches should fall on the solid line. All others will be left with a constant giving-up density $GUD = C/A$. However, all the poor patches will be left at one of the stopping points along the solid line in Fig. 4A. For these, GUD should hence be strongly negatively related to search time.

For less extreme distributions (rich to poor ratio 2 or slightly less) the relation between GUD and search time will be less clear, and often not monotonic (Fig. 4). For the distribution with rich-to-poor ratio two (Fig. 4B), 83% of the poor patches, and 10% of the rich patches, will fall on the negative relation. Most of the remainder of the patches will have GUDs equal to C/A. Thus, in this case too, it should be possible to identify the negative correlation between GUDs and search time.

For distributions with smaller differences between rich and poor patches (rich to poor ratio 1.5 or less; Fig. 4C-D) the optimal Bayesian strategy will effectively be one with constant quitting harvest rates.

The explanations for these different patterns are the same as in the smooth distributions (Olsson and Brown 2006). When the difference between rich and poor patches is small all the information the forager gains from finding a prey item is essentially negative. The patch now likely contains less than it did before finding the item (Iwasa et al. 1981). Hence, there is no FBI.

When rich patches are much more valuable than poor, finding an item will often increase the estimate of patch

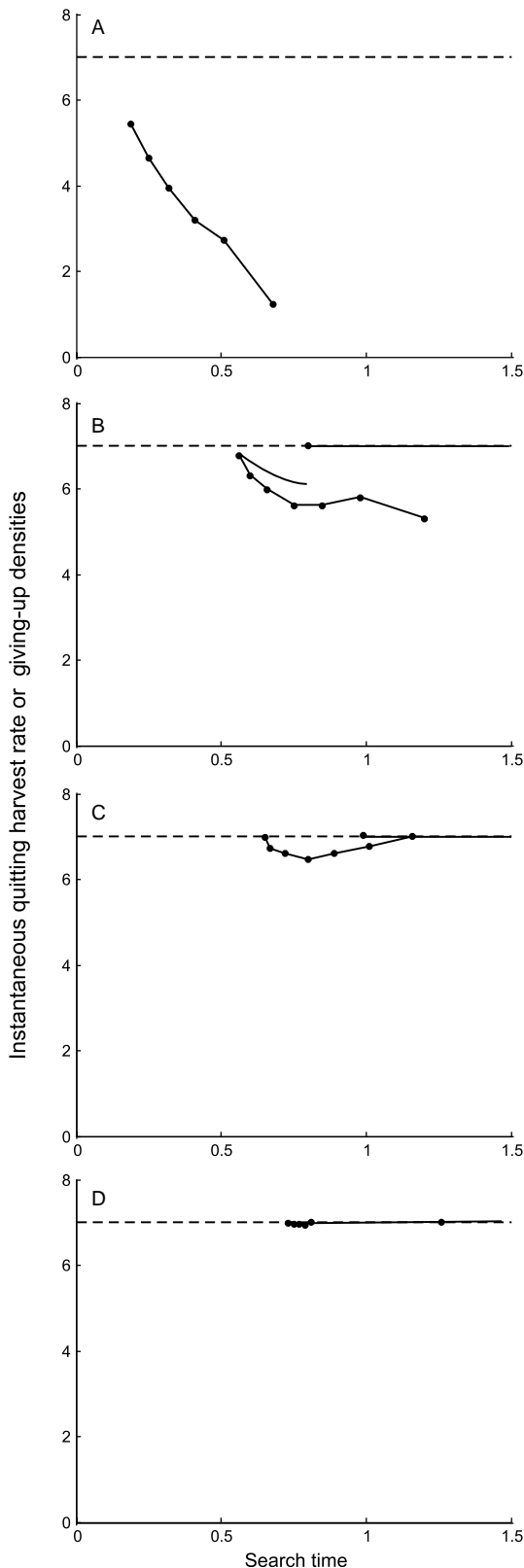Instantaneous quitting harvest rate or giving-up densities

Search time

Fig. 4

quality. There is good news to be had, and FBI is positive.

These are properties of the two-patch distributions that are similar to the binomial (no FBI) and negative binomial distributions (positive FBI: Olsson and Brown 2006). However, there is also one striking difference. In a negative binomial distribution the quitting harvest rate increases with search time (Olsson and Holmgren 1998), i.e. FBI decreases. This is mostly because when newly arrived in a patch sampling information is valuable, and therefore a low instantaneous intake rate is accepted to gain information about patch quality (Olsson and Holmgren 1998, Dall et al. 2005, Olsson and Brown 2006). By contrast, in a two-patch distribution, the forager generally accepts a lower instantaneous intake rate, the more prey it has found. I suggest that is because it is getting closer to the information state where it will switch estimate of patch quality, from likely poor to surely rich (Fig. 1). A forager exploiting a negative binomial distribution gets a more robust estimate with the number of prey found. Such a forager that has already spent a long time in a patch, and hence found many prey, will not be much impressed by yet one item. One that has exploited a patch in a two-patch system for a long time will dramatically change its estimate of patch quality if it finds one more item (Fig. 1).

Giving-up densities

GUDs can be very useful for evaluating patch leaving strategies in general, including Bayesian foraging (Brown 1988, Valone and Brown 1989, Meyer and Valone 1999, Olsson et al. 1999, van Gils et al. 2003, Vásquez et al. 2006). The frequency distributions of GUDs left in rich and poor patches depend heavily on the proportion of rich and poor patches in the environment (Fig. 5).

When the poor patch is the more common (Fig. 5C) the leaving policy is adjusted to exploit poor patches effectively. This is done by shifting the stopping points to the left (Fig. 2B, line 3). The forager adopts a rather hasty policy, leaving most patches when hitting the lower horizontal lines in the stopping points. Poor patches will be rather correctly utilized, but rich patches will be heavily under-used. The distribution of GUDs is strongly weighted to the region close to the C/A-line, for poor patches. For rich patches GUDs are bimodal.

The difference in use between the two patch types is shown also by Fig. 7. It presents the proportion of prey harvested from the rich patch, as a function of the

Fig. 4. Predicted relations between total search time spent in a patch and the instantaneous quitting harvest rate, at which it is left. The simulations shown are 8 (A), 1 (B), 10 (C) and 11 (D). The dashed horizontal line corresponds to the respective upper horizontal lines in Fig. 2 and 3, i.e. when $n = N_2 - C/A$. The solid lines and curves correspond to the stopping points, and lower horizontal lines in Fig. 2 and 3.

Fig. 5. Frequency distributions of GUDs of poor and rich patches. The width of the horizontal bars indicate the relative frequency of patches predicted for a certain GUD. The dots indicate the average GUD. The panels are for simulations 1 through 6 (panels A through F).
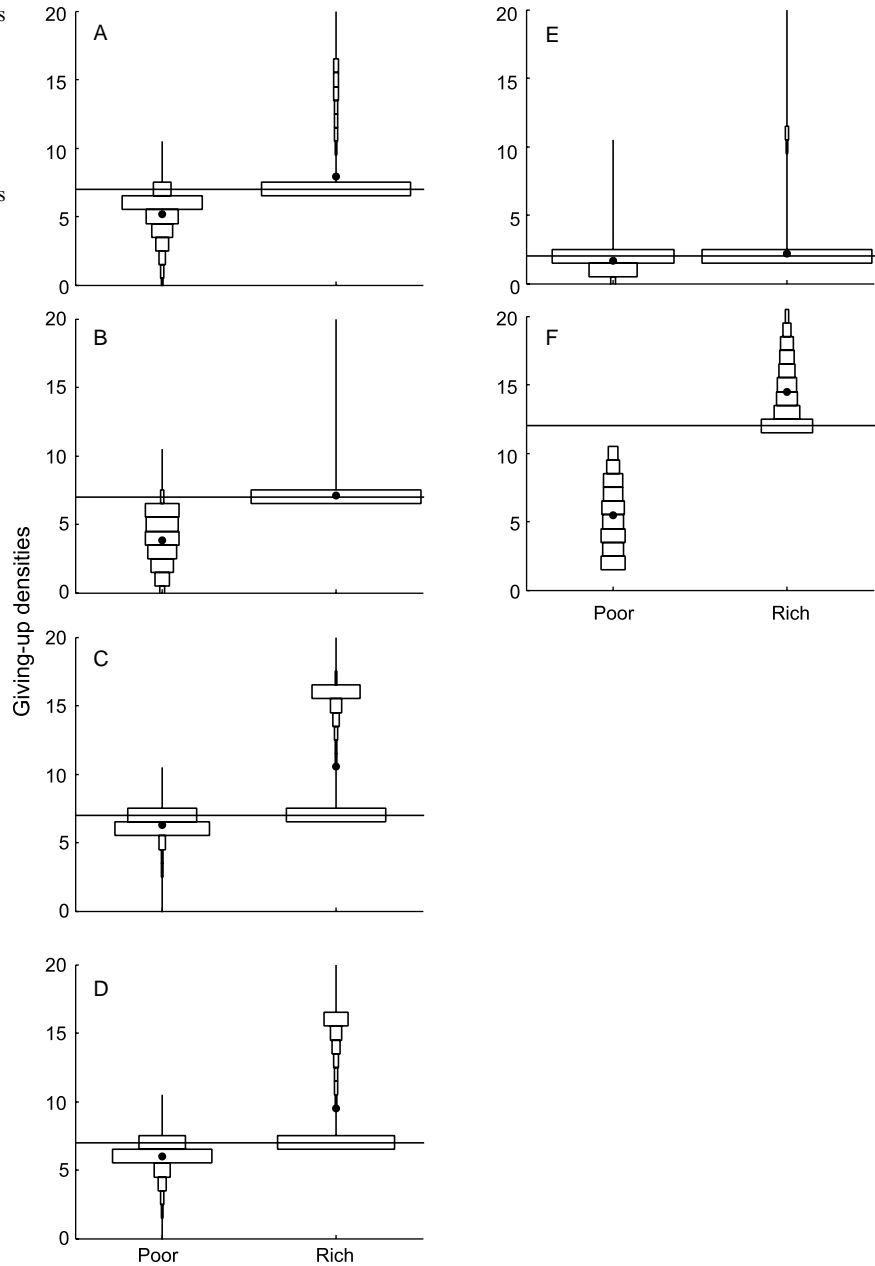
proportion harvested from the poor. The figure also indicates the situation when GUD = C/A, which is what a perfectly informed forager would do, and the 1:1 line, where a perfectly ignorant forager's points would be. A perfectly informed forager would leave all patches at a constant quitting harvest rate, and consequently at a constant GUD (Brown 1988). Therefore, the predictions for a perfectly informed forager would generally be that the difference between GUDs in rich and poor patches is larger than for a Bayesian forager. That is, their points

should appear further from the 1:1 line. As a contrast, the completely ignorant forager spends the same amount of time in all patches, and therefore leaves rich and poor patches after having harvested the same proportion from each. A forager with a fixed number strategy would have its points along a line rooted in the origin with slope N_1/N_2 in Fig. 7, i.e. always quite far below the 1:1 line.

A forager exploiting a two patch system with pre-dominantly rich patches adopts a policy that much ignores the poor ones (Fig. 5B, Fig. 7 point 2). The
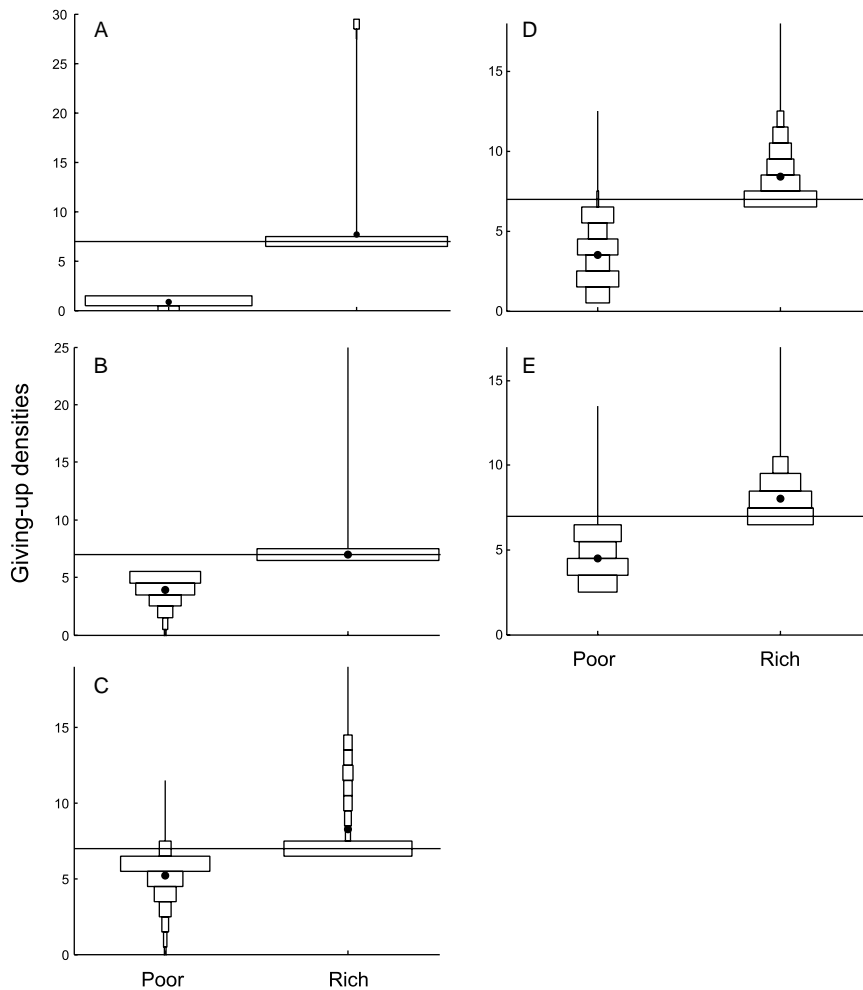
Fig. 6. As in Fig. 5, but for simulations 7 through 11 (panels A through E).

consequence is that rich patches are well used, but poor patches are over-used. Rich patches have GUDs very close to C/A, but poor patches have a wide distribution of GUDs.

The frequency distribution of GUDs also depends on the critical intake rate C chosen by the forager (compare Fig. 5A with 5E and 5F). With a low C all patches are harvested almost to exhaustion. The forager will then have very good information about patch quality, and use both rich and poor patches almost like a perfectly informed forager (Fig. 7 point 5). Most patches will have GUD ≈ C/A. When C is high (Fig. 5F) all patches must be left soon, to try to gain the high rate (C). In fact, a perfectly informed forager would leave poor patches without searching them at all (Fig. 7 point 6). The resulting frequency distribution of GUDs has a very high variance, and neither poor nor rich patches are very well used.

In fact, in Fig. 5F, C is so high that the forager will never gain perfect information ($N_2 - C/A < N_1$; Fig. 1B). In such cases a higher proportion of prey are harvested

from the poor patch, than from the rich. This is shown by the point appearing below the 1:1 line in Fig. 7 (point 6). Also in simulations 10 and 11, $N_2 - C/A < N_1$. However, here C/A is clearly lower than N_1, but instead the difference in quality between the rich and the poor patch is small. Again, the points appear below the 1:1 line in Fig. 7, but in these cases not so far from where a perfectly informed forager would put them.

Figures 3 and 6 show the set of cases (B) where the frequencies of rich and poor patches are equal, and rich and poor patches are symmetrically distributed around 15. When the difference between rich and poor patches is large (Fig. 3A, 6A-B), a moderate sampling effort will give the forager a very reliable (or even perfect) estimate of patch quality. The optimal policy allows it to use the rich patches very close to perfectly. It does so by accepting quite low instantaneous intake rates early on during the stay in the patch. That way it is guaranteed to correctly identify virtually every rich patch. But it also means that the forager will often take several prey items from the poor patches, even though the "prescient ideal"
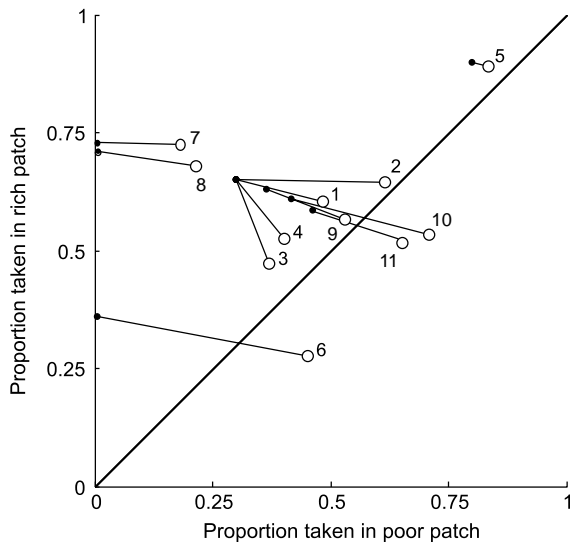
Fig. 7. The proportion of prey taken from rich (vertical axis) and poor (horizontal axis) patches. The open circles indicate the result of the optimal Bayesian strategies. The black dots indicate the prediction for a prescient forager, exploiting the same environment. These points are connected by lines. The diagonal line (1:1) indicates the situation where equal proportions are harvested from both rich and poor patches, i.e. the predicted outcome of fixed time foraging. The open circles appearing above the 1:1 line are for cases when Bayesian foragers gain perfect information before departure.

would leave all poor patches instantly (as $N_1 < C/A$ in this distribution). Thus, the price for finding the rich patches is a quite heavy overuse of the poor patches.

Figure 6A gives a very useful prediction that can be used in experiments. In this case (simulation 7) the poor patch only contains 1 item, and the rich 29, and the quitting harvest threshold $C = 7$. The optimal policy is to leave the patch after 0.12 time units if no prey have been found or after 0.25 time units if only one has been found by then. If more than one item has been found before that time, the patch is searched until 23 items have been found. With such a policy most poor patches are left without harvesting any items, and most rich patches are left with a GUD of 7. However, the signature that reveals a Bayesian forager is in how many poor patches are left after having taken the single available item, or how many rich patches are left without taking any. The expectation from the current model is that in approximately 11% of the poor patches the prey item should be removed, and in approximately 3% of the rich patches no items should be taken.

When the difference between rich and poor patches is small (Fig. 3C–D, Fig. 6D–E), the rich patches will often be underused. The reasons are threefold. It is more difficult to discriminate between similar patches. It pays less to discriminate between them. The forager never acquires near perfect information unless $N_2 - N_1 > C/A$.

Discussion

There have been several empirical studies demonstrating that animals are capable of Bayesian updating (reviewed by Valone 2006). However, a species evolved to use Bayesian foraging in one context (e.g. patches with prey in a negative binomial distribution: Olsson et al. 1999, van Gils et al. 2003), is not necessarily capable of using the optimal Bayesian rules for a different context, such as a two-patch system. However, Vásquez et al. (2006) provides a nice demonstration that two-patch systems seem to be well handled by real animals, even if their natural patches are of different kinds.

Strict two-patch systems may not be common in nature. It may be more common to find cases where two patch types differ in their mean rewards, but where there is a variation around each mean. If the two distributions do not overlap much, the general predictions and conclusions from this study should still hold.

The great advantage of two-patch systems is obviously that they are easy to use in an experimental context. Creating an experimental environment with a realistic negative binomial distribution is a tremendous endeavor that has been completed only once (van Gils et al. 2003). The difficulty lies in the mere number of patches that the experimenter needs to provide more or less simultaneously. Two-patch systems are much easier to create, with a minimum of only two patches (Meyer and Valone 1999, Vásquez et al. 2006).

It is worth pointing out, that I have modeled distributions of two different kinds (set A and B, see beginning of Results for explanation), and as far as I know only set B has been used in experiments. Set B, where $Pr_1 = Pr_2 = .5$, is ideal for experiments with pair wise patches. A pair of close patches in the same microhabitat allows the assumption that both patches should have been treated identically, if they had the same food abundance. This makes e.g. GUD analyses straight-forward (Valone and Brown 1989, Valone 1991, Meyer and Valone 1999, Vásquez et al. 2006). The alternative way of creating a two-patch system with $Pr_1 \neq Pr_2$ (set A) may make pairs of patches less useful. Also, it may require a higher number of patches. If the frequencies of the patches are, say, 0.1 and 0.9 respectively, then at least 10 patches are needed, and so on. However, using set A may sometimes increase the flexibility of two-patch systems, and also generates certain specific predictions.

I have previously (Olsson and Holmgren 1998, 2000) criticized the usage of the terms overuse and underuse. In this paper I have been using them frequently. The reconciliation lies in the possibility to find the "template usage" to which the Bayesian usage should be compared. When dealing with the smooth distributions, and e.g. a rate maximizing policy, this is not straightforward. It may even be confusing and misleading. When dealing with a two-patch distribution and a fixed critical

potential rate, it is much more straightforward and the terms may be useful (Valone and Brown 1989, Meyer and Valone 1999).

Testable predictions

The model generates at least three interesting qualitative predictions, which should be possible to test in experiments.

The bimodal distribution of GUDs of rich patches in Fig. 5C and 5D, is predicted only under a Bayesian policy, as described above. It occurs mainly, at least is strongest, in cases when rich patches are rare, i.e. $Pr_2 <$ 0.5. Neither a prescient (perfectly informed) forager, nor a fixed-time (ignorant) forager would exploit patches like this. The prescient would ideally leave all patches when GUD $= C/A$, i.e. generate zero variance in GUD. Some random error around this expectation should probably be tolerated in an empirical study, but that should not generate a bimodal distribution. A fixed time forager is predicted to generate a smooth, unimodal distribution of GUDs, with the proportion of prey taken in the rich patch equal to that of the poor. This can be seen by substituting Δ in Eq. 5 with any fixed time to spend in all patches. Both the prescient and the fixed-time forager should generate the same frequency distribution of GUDs, regardless of the frequency of rich patches in the environment. That is, their distributions of GUDs should only depend on C.

Another prediction is that the ability to gain perfect information depends on the rate that the foragers demand to remain in the patch, C. When C is high information about patch quality is always incomplete (simulation 6). This critical rate may be higher e.g. in micro-habitats with higher predation risk, or when the missed opportunity cost of foraging is high (Brown 1988, Brown and Alkon 1990, Olsson et al. 2000). Thus, the overuse of poor patches and underuse of rich patches is expected to be stronger in situations when the forager must gain much energy to remain in a patch (Schmidt and Brown 1996, Meyer and Valone 1999). The proportion of prey harvested in the rich patches should be lower in a risky habitat, than in a safe (compare points 1, 5, and 6 in Fig. 7), for a Bayesian forager. This is directly testable by providing pairs of rich and poor patches simultaneously in high and low risk microhabitats. Meyer and Valone (1999) present data that seems to agree with this prediction. However, in their study C is perhaps not constant across alternatives, as they compare different periods in time.

Let us return to simulation 7, in which 11% of the poor patches had its single prey item removed. If predation risk, or some other foraging cost decreased then the proportion of empty poor patches would increase. If C was as low as one, then this proportion

would be as high as 30%. In contrast, practically no rich patches would be left un-harvested.

Finally, the negative relation between search time and GUD (Fig. 4) is only predicted for Bayesian foragers exploiting a two-patch distribution, and for fixed-number foragers. Obviously, a fixed number forager should leave all patches of a given initial quality with the same GUD. A Bayesian forager should not, unless the quitting potential, C, is very low. Importantly, a fixed number strategy should not be expected in a two-patch system, as it is not optimal under any reasonable assumptions. Only a forager capable of counting accurately, but not keeping track of search time should use such a strategy. Thus, only if one wants to ponder over the curious actions of stupid animals (Milne 1926, Adler and Kotar 1999) need such a strategy be considered.

The negative relation between GUD and search time should be most pronounced when the difference between the rich and poor patch is large. The non-linear relation predicted for intermediate rich-to-poor ratios, will likely often appear as a weak negative relation. In both these cases all, or most, of the rich patches, and some of the poor patches, will be left with the GUDs along the horizontal line in Fig. 4. In two-patch systems with low rich-to-poor ratios, GUDs are expected to be effectively independent of search times.

Collecting data on search times, in addition to the GUD data, may be very useful, also in a two-patch system (Olsson et al. 1999, Olsson and Holmgren 1999, van Gils et al. 2003). In a foraging experiment e.g. video analysis of search times may require some additional effort, which may sometimes not be feasible (Vásquez et al. 2006). As an alternative to such labor-intensive methods, it is possible to use transponder systems (Olsson et al. 2001, 2002). However, this requires that patches be predominantly used for foraging, and not for other activities.

I hope that the predictions generated in this study will inspire new experiments, which may improve our understanding of animals' information use.

References

Adler, F. R. and Kotar, M. 1999. Departure time versus departure rate: how to forage optimally when you are stupid? – Evol. Ecol. Res. 1: 411–421.

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. and Alkon, P. U. 1990. Testing values of crested porcupine habitats by experimental food patches. – Oecologia 83: 512–518.

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

Dall, S. R. X., Giraldeau, L. A., Olsson, O. et al. 2005. Information and its use by animals in evolutionary ecology. – Trends Ecol. Evol. 20: 187–193.

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. 1987. Stochastic models of optimal foraging. – In: Kamil, A. C., Krebs, J. R. and Pulliam, H. R. (eds), Foraging behavior. Plenum Press, pp. 273–302.

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

Green, R. F. 2006. A simpler, more general method of finding the optimal foraging strategy for Bayesian birds. – Oikos.

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.

Lima, S. L. 1984. Downy woodpecker foraging behavior: efficient sampling in simple stochastic environments. – Ecology 65: 166–174.

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

Meyer, M. D. and Valone, T. J. 1999. Foraging under multiple costs: the importance of predation, energetic, and assessment error costs to a desert forager. – Oikos 87: 571–579.

Milne, A. A. 1926. Winnie the Pooh. – Methuen & Co. Ltd.

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

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.

Olsson, O., Wiktander, U. and Nilsson, S. G. 2000. Daily foraging routines of a small bird feeding on a predictable resource. – Proc. R. Soc. B 267: 1457–1461.

Olsson, O., Brown, J. S. and Smith, H. G. 2001. Gain curves in depletable food patches: a test of five models with European starlings. – Evol. Ecol. Res. 3: 285–310.

Olsson, O., Brown, J. S. and Smith, H. G. 2002. Long- and short-term state-dependent foraging under predation risk: an indication of habitat quality. – Anim. Behav. 63: 981–989.

Schmidt, K. A. and Brown, J. S. 1996. Patch assessment in fox squirrels: the role of resource density, patch size, and patch boundaries. – Am. Nat. 147: 360–380.

Stephens, D. W. and Krebs, J. R. 1986. Foraging theory. – Princeton Univ. Press.

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

Valone, T. J. 1992. Information for patch assessment: a field investigation with black-chinned hummingbirds. – Behav. Ecol. 3: 211–222.

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.

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.

Vásquez, R. A., Grossi, B. and Márquez, I. N. 2006. On the value of information: studying changes in patch assessment abilities through learning. – Oikos 112: 298–310.

Subject Editor: Per Lundberg