FORAGING AND FLOCKING STRATEGIES: INFORMATION IN AN UNCERTAIN ENVIRONMENT

COLIN W. CLARK AND MARC MANGEL

Department of Mathematics, University of British Columbia, Vancouver V6T 1W5, Canada; Department of Mathematics, University of California, Davis, California 95616

Submitted December 16, 1982; Accepted October 6, 1983

In order to survive and pass on its genetic material, an animal must eat and avoid being eaten. Many patterns of behavior have apparently evolved in response to these basic needs (Krebs and Davies [1978], especially chap. 1, 2). Here we study questions concerning the behavior of animals searching for food in a "patchy" environment. Our models pertain both to the behavior of individual animals and to the behavior of groups of animals.

Given that the distribution of forage is patchy, the role of information pertaining to the location, size, and quality of food patches becomes significant. At one extreme, the forager would know these details completely, but would still face the problem of efficient utilization of forage patches. This problem has been investigated theoretically under the aegis of (deterministic) optimal foraging theory (Charnov 1976; Pyke et al. 1977; Cowie and Krebs 1979; Krebs 1978; McNair 1982).

At the opposite extreme, nothing would be known initially about the location and abundance of forage patches, and the forager could only obtain this information via sampling. More realistically, at any given time the forager will possess partial information about its forage supply. For example, having spent a certain amount of time searching for prey in some traditional feeding area, a predator may be able to compare its observed encounter rate with the "usual," or prior expected catch rate for the given area, or other areas. If the rate is much lower than normal, the predator may decide to switch to some other area.

Searching for food items thus has both a short-term benefit (food items are located and eaten), and a long-term benefit (information is accumulated which may be valuable in improving future foraging success). An optimal foraging strategy must somehow involve a balance between these short-term and long-term benefits. The precise nature of optimal strategies seems far from obvious.

We shall systematically investigate the role of information in foraging strategy, using a number of simple models. The possibility of group foraging is discussed in the first two sections. In the third section a Bayesian approach is used to indicate how the individual forager may update its estimate of patch quality. Finally,

Am. Nat. 1984. Vol. 123, pp. 626–641.

returning to the case of group foraging, we consider the effect of group size on the variance in individual feeding rates. Some of our results overlap with those of Caraco (1981).

A GAME-THEORETIC INTERPRETATION OF FLOCKING

We first discuss a preliminary, simplified model of flocking strategy, in which the advantages of flocking stem entirely from the increased amount of information obtained by individual birds. (To be explicit, we shall henceforth refer to bird flocks, but our models apply equally to other animals.) For this model, flocking is an "evolutionarily stable strategy" (ESS), even though it does not necessarily lead to an overall increase in feeding rate for the entire flock. A more complex model to be introduced later will result in important modifications of this preliminary conclusion.

Consider an individual bird, searching over a given area for food that occurs in randomly distributed patches of uniform size. To keep the analysis simple, let us ignore the depletion of patches. Under these assumptions the probability of encountering a patch of forage in an infinitesimal time interval $(t,t + dt)$ is independent of the foraging pattern used, and is given by

$$p = \lambda dt \tag{1}$$

where λ is a constant equal to the average rate of encountering patches. If the forager finds at most one patch in dt and searches randomly, then equation (1) leads to a Poisson process. Alternatively, one can simply assume that the forager encounters patches according to a Poisson process with parameter λ. (The Poisson model can be modified to cover clumped distribution of patches [Clark 1983], nonuniform patch size [Mangel and Clark 1983], and depletion of patches [see Appendix]).

Let τ denote the time required for one bird to eat one patch; the bird eats the entire patch, and then resumes searching. Since the expected time to encounter a patch equals $1/\lambda$, the average long-term feeding rate is

$$f_\lambda = \frac{B}{\tau + 1/\lambda} = \frac{B\lambda}{\tau\lambda + 1} \tag{2}$$

where B denotes the weight of one patch (see Appendix).

Suppose next that two birds are foraging over the same area. Each bird can choose one of two strategies: N (nonflocking strategy): the bird searches and feeds independently of the other bird; or F (flocking strategy): while searching, it also observes the other bird, and intrudes whenever the other bird finds a patch. In this way the bird doubles the amount of information it receives about patch location. We calculate the average feeding rate for each bird, depending on the strategies used by both, assuming that the two birds divide the food in the patch equally, and that no time is spent fighting over the patch.

If both birds use the nonflocking strategy N, both feed at the rate given by equation (2). If both birds use the flocking strategy F and the time to reach the other bird's patch is small compared to τ and $1/\lambda$, then the expected time between

encountering patches is approximately $1/2\lambda$, the time to each a patch is approximately $\tau/2$, and each bird eats $B/2$ per patch. Thus, to first order, the average feeding rate for each bird stays the same, as in equation (2). Finally, if only the first bird uses the flocking strategy, it can be seen that the feeding rates for the two birds are (to first order) $2f_\lambda - f_{\lambda/2}$ and $f_{\lambda/2}$, respectively. (The second bird gets to eat only half the patches it finds, so that its expected feeding rate is $(B/2) \div (1/\lambda + \tau/2) = f_{\lambda/2}$, etc.). (We assume here that the first bird intrudes on the second whenever the second bird locates a patch, even if the first is eating a patch of its own when this occurs. The first bird is then able to relocate its own unfinished patch.)

These results lead to the following individual "payoff" matrix for the flocking "game":

	bird 2	bird 2
bird 1	N	F
N	f_λ	$f_{\lambda/2}$
F	$2f_\lambda - f_{\lambda/2}$	f_λ

where the entries represent the expected payoffs in terms of average feeding rates, for bird 1. Since $2f_\lambda - f_{\lambda/2} > f_\lambda$ it follows that the flocking strategy (F) is a Nash competitive equilibrium (Nash 1951) in the sense that no matter which strategy is employed by bird 2, the first bird increases its payoff by using the flocking strategy. The Nash equilibrium is also an evolutionarily stable strategy (ESS) in the sense of Maynard Smith (1974)—provided, as is usually taken for granted in foraging theory, that increased foraging rates bestow increased "fitness" upon individuals (see e.g., Krebs and Davies 1978, p. 23).

In the foregoing model there is no overall advantage to flocking, the diagonal matrix entries being equal. Other considerations, such as conflict or aggression, could lead to a payoff matrix in which the (F,F) strategy is actually inferior to (N,N), for example,

	N	F
N	10	5
F	13	8.

In this situation, flocking would still be the Nash equilibrium (and ESS), yet each bird would lose if both used the flocking strategy. (This game has the well-known structure of the "prisoners' dilemma," which is also the basis of Garrett Hardin's "tragedy of the commons" [Hardin 1968].)

The following objection to the above argument might be raised at this point. The bird that feeds alone experiences a feeding rate f_λ. If it intrudes on another bird, it knows from experience that the other bird will also intrude on it, and both birds will thus feed at the same rate f_λ, or possibly at an even lower rate. Hence there is no advantage in flocking. Arguments of this kind—strategy versus answering strategy—arise throughout game theory, and are not easily dealt with. Our next, more realistic model of flocking strategy, however, shows that under certain conditions both birds stand to benefit from flocking together.

A MORE COMPLEX MODEL

We now introduce a more complex and realistic model of flocking behavior. The model leads to a correspondingly richer set of predictions.

Assume that n birds are searching for food over a given area. As before, the food occurs in randomly distributed patches of equal, known size.

Suppose that each bird has a capacity C, per "meal," in the sense that the bird ceases feeding if it obtains C from a given patch. Let $B = mC$, i.e., each patch contains exactly m "meals." Let τ_1 denote the feeding time required to eat one meal.

Under the flocking strategy all n birds search independently for patches of food, simultaneously observing each other. When a patch of food is located, all n birds converge and feed on the patch until they are either all satiated, or the patch is exhausted. Satiation occurs first if $m > n$, whereas patch exhaustion occurs first for $m < n$. (Satiation would presumably also affect subsequent searching behavior, but we shall ignore this complication here.) Clearly the model could be further elaborated in many ways; later we shall include crowding effects.

To calculate the expected feeding rate, note first that the amount taken by each bird from a given patch is θC, where

$$\theta = \min(1, m/n).$$

Also, the time required to eat a patch is $\theta\tau_1$, and the expected search time is $1/\lambda n$. Thus the expected feeding rate is given by

$$f_\lambda(n) = \frac{\theta C}{\theta\tau_1 + 1/\lambda n} = \frac{C}{\tau_1 + 1/\lambda \min(m,n)}. \tag{3}$$

This is shown in figure 1 (upper curve). Since patches are encountered more frequently when information is shared among n birds, the average feeding rate increases up to $n = m$, the number of birds required to consume an entire patch. Beyond this size, flocking produces no overall increase in the rate of feeding since any increase in the rate of finding patches is compensated by a decrease in the size of the meal eaten (if $n > m$, only m/n of a full meal is eaten).

Consider next the individual bird's choice of strategy: N (nonflocking) versus F_n (join an existing flock of n birds). For nonflocking, the expected feeding rate is $f_\lambda(1)$ (assuming that the rest of the flock does not intrude on the individual bird; if intrusion does occur, the rate is less than $f_\lambda(1)$, and the feeding rate for flocking is $f_\lambda(n + 1)$.

Thus an individual bird always does better by joining an existing flock than by feeding alone. Similarly, any flock of size less than m is unstable, since its members will do better by joining a larger flock. The model thus predicts that arbitrarily large flocks will form. However, flocks of size greater than $2m$ are only "neutrally stable," in the sense that if by chance they break up into smaller flocks of size $\geqslant m$, there will be no tendency for them to recombine. Other possible advantages of large flock size, such as protection against predation (e.g., Pulliam et al. 1982), are not under consideration here.

Clearly, however, various second-order effects will tend to reduce the feeding

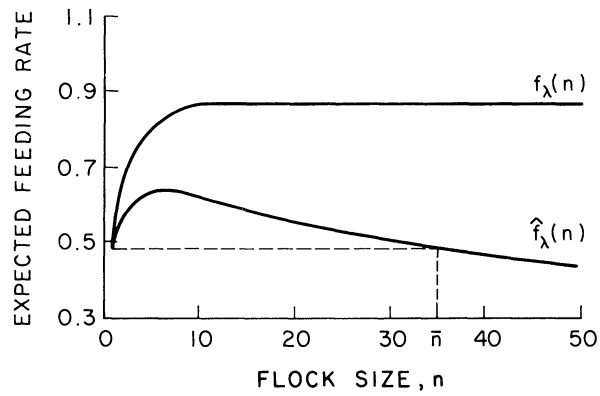

FIG. 1.—Expected individual feeding rates for birds in flocks of size n; upper curve $f_\lambda(n)$ ignores "second-order" effects, which are included in the lower curve $\hat{f}_\lambda(n)$ (see Appendix). The optimal flock size is $\bar{m} = 7$, but the natural equilibrium is $\bar{n} = 35$. (Parameter values: $\lambda = 1$, $\tau_1 = 1$, $m = 10$, $T = 25$, $\epsilon = 0.15$. Curves are drawn continuously for convenience, even though actually defined only for integral values.)

rate of individual birds in large flocks. Patches located in one part of the flock will not be shared by distant birds (i.e., information will not be shared throughout the whole flock), conflicts between neighboring birds will increase with flock density, and so on. We shall model such crowding effects in an ad hoc fashion by introducing a crowding factor ϵ_n, so that the net average feeding rate now becomes

$$\hat{f}_\lambda(n) = \epsilon_n f_\lambda(n). \tag{4}$$

Crowding is assumed to become progressively more deleterious to feeding as flock size increases: $1 = \epsilon_1 > \epsilon_2 > \epsilon_3 > \ldots$. The function $\hat{f}_\lambda(n)$, computed from a specific model of crowding (see Appendix), is shown as the lower curve in figure 1.

The modified model implies the existence of a socially optimal flock size $\bar{m}$, for which the average individual feeding rate is maximized. This optimal flock size is less than or equal to m, the number of meals per patch.

What size flocks will actually be observed? Assume that birds make their decisions independently, joining existing flocks or changing flocks so as to maximize their own feeding rate. Suppose first that only one flock exists, of size $n < \bar{m}$ (where $\bar{m} > 1$). A new arrival will then join the flock, thereby increasing the feeding rate of all individuals. But even when $n \geqslant \bar{m}$, a new arrival will join the flock, rather than feeding alone, as long as $n < \bar{n}$, where (see fig. 1)

$$\hat{f}_\lambda(\bar{n}) = \hat{f}_\lambda(1). \tag{5}$$

Clearly $\bar{n} > \bar{m}$ unless $\bar{m} = 1$ (this possibility is discussed later). Since $\bar{n}$ is an integer, it is the integer closest to the solution of (5), with $\hat{f}_\lambda(\bar{n}) \geqslant \hat{f}_\lambda(1)$, and $\bar{n} \geqslant 1$.

Note that the members of a flock of size $\bar{n}$ could all increase their feeding rates if they could divide into, say, two flocks of size $\bar{n}/2$. No individual bird has the incentive to break away alone, since $\hat{f}_\lambda(\bar{n} - 1) > \hat{f}_\lambda(1)$. If a few birds did break

away together, however, our hypothesis implies that two flocks of equal size $\bar{n}/2$ would then form. With $\hat{f}_\lambda(n)$ having the shape in figure 1, we still have $\bar{n}/2 > \bar{m}$, and in fact four flocks of size $\bar{n}/4$ would be even better than two. The penalty for a single bird breaking away from a flock of size $\bar{n}/2$ is large, however, so these flocks should be highly stable, even though suboptimal.

Clearly an equilibrium can occur only when birds are distributed in flocks of size $n_1, n_2, \ldots$ in such a way that $\hat{f}_\lambda(n_i)$ is the same for all i (as closely as possible). Also we must have $n_i \geqslant \bar{m}$ for otherwise birds can increase their own feeding rate by switching flocks. Thus all flocks will be as close as possible (± 1) to equal size.

It is worthwhile to work out the successive distributions of flock sizes as n, the total number of birds, increases. For example, let $\bar{n} = 35$, as in figure 1. Then for $n \leqslant 35$ a single flock forms; for $n = 36$, two flocks of size 18 form; and so on. In general, if $n = 35k$, a multiple of $\bar{n}$, then k flocks of size 35 will form. For $n = 35k + 1$, sufficiently many birds leave the k flocks to form $k + 1$ flocks of nearly equal size. For example, if $k = 33$ and $n = 35k + 1$, then 34 flocks of size 34 birds each will form. For n larger than this, all flocks will be of size 34 or 35.

To summarize, flocking of foraging birds, even in the absence of predators, in an equilibrium strategy (and hence, under the assumptions noted previously, an evolutionarily stable strategy), provided that food is both scarce and patchy. We note that the equilibrium flock is in general larger than the socially optimal flock. In fact, we have shown that the socially optimal flock is not evolutionarily stable.

Our model provides several predictions that are both nontrivial and testable. Birds in foraging flocks will constantly observe one another's activities, and will share food patches when they are discovered. Mixed-species flocks will have common forage requirements. The average feeding rates for different flocks in a given area will be approximately equal. Flock sizes will be larger than the optimum, and in particular larger than the average number of "meals" per patch.

The model can also be used to test the hypothesis that, in specific cases, flocking is primarily associated with foraging in a patchy environment, rather than with defense against predation. If flocking is associated with foraging, our model predicts that the size of flocks will respond both to the overall abundance of food, and to the patchiness of its distribution (as elaborated below); this would not be the case for antipredator flocking. For example, Cody (1971) presents evidence that the size of finch flocks in the Mohave desert is responsive to food abundance; no instances of predation were observed on these flocks.

The foraging behavior of lions in the Serengeti has been analyzed by Caraco and Wolf (1975), who found that observed sizes of foraging groups often exceeded the optimum size in terms of feeding rate. Lions feeding on wildebeest, for example, formed groups of 5 to 7 lions, compared with the optimum of 2 lions per group. On the other hand, when feeding on Thomson's gazelle, which provides only about 2 lion "meals" per gazelle, average groups were from 1.5 to 2 lions. Both of these observations are in close accord with the value $\bar{n}$ predicted by our analysis, but Caraco and Wolf suggested that the supra-optimal group sizes could be attributed to other ecological factors such as reproductive success, competition with hyenas, and so on.

Let us elaborate on the relationship between abundance, patchiness, and flock

size. First note that for large λ we have from equation (3)

$$f_\lambda(n) \approx C/\tau_1 \qquad \text{for all } n. \tag{6}$$

In this situation birds spend most of their time feeding, at the given rate C/τ_1. Second-order effects of flocking will dominate, rendering $\hat{f}_\lambda(n)$ a decreasing function for $n \geqslant 1$. Thus flocks will not form when food patches are abundant, and in general the size of flocks will increase in proportion to the scarcity of food patches.

Next we consider the effect of patchiness. Keeping the overall food abundance $B\lambda$ constant, we compare the cases of large, scarce food patches (B large, λ small) versus small, abundant patches (B small, λ large). In the first case, the number of meals per patch, m, is large; also by equation (3) we have $f_\lambda(n) \approx nC\lambda$ for $n \leqslant m$. In the second case, m is small and $f_\lambda(n) \approx C/\tau_1$ for $n \leqslant m$. The resulting curves f_λ and $\hat{f}_\lambda$ are shown in figure 2. In the second case (low patchiness), flocking is suppressed.

We conclude that flocking is advantageous precisely in those situations in which it yields valuable information. When food is abundant, information as to location is not very valuable; if food is scarce but evenly distributed, little information is derived from searching. Flocking is most valuable when food is both scarce and patchy.

For species which are subject to predation, our model yields a further prediction, resulting from the fact that the evolutionarily stable aggregation strategies for foraging and for antipredation are different. The antipredation strategy, when the group relies upon sentries (rather than group defense), calls for large, closely packed aggregations, whereas the foraging strategy yields aggregations of a limited size, with sufficient individual spacing to allow for fairly independent search. One therefore should expect to observe conglomeration and increased density of foraging aggregations in the presence of predators. Such behavior has often been reported.

INFORMATION AND OPTIMAL FORAGING

The foraging behavior of an individual predator seeking food in a patchy environment has been modeled using a class of theoretical results called marginal value theorems (Charnov 1976). Experimental results have also been reported (Pyke et al. 1977; Cowie and Krebs 1979) with varying degrees of qualitative agreement with the theoretical predictions.

The marginal value theorems address the question, When should a forager leave the patch in which it is feeding and search for a new patch? Charnov's original formulation of the marginal value theorem stated that the predator should leave the patch when the instantaneous feeding rate within the patch declines to the maximum average feeding rate over the available foraging area.

As emphasized by Oaten (1977), however, foraging is an inherently stochastic process. Attempts to determine optimal foraging strategies by means of deterministic models, such as that of Charnov (1976), may be seriously misleading. Oaten develops a stochastic foraging model in which (1) prey encounters within a patch

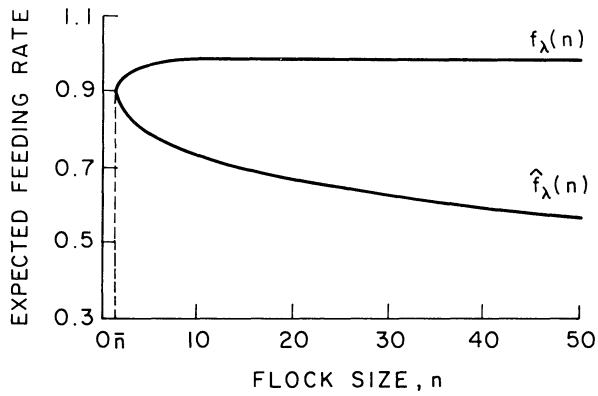

FIG. 2.—Expected individual feeding rates for birds in flocks of size n. Here λ, proportional to the number of patches, is large, and flocking is suppressed. (Parameter values: $\lambda = 10$, $\tau_1 = 1$, $m = 10$, $T = 25$, $\epsilon = 0.10$.)

are random, and (2) the number of prey initially in a given patch is a random variable whose distribution is known to the predator, and is the same for all patches.

Two benefits are provided by foraging in a given patch: food is located, and information is obtained about the likely quality (expected long-term feeding rate) within the patch. This information can be used by the forager in deciding when to move to another patch.

The informational aspects of the foraging problem can be studied by using Bayesian methods (Green 1980; Iwasa et al. 1981). To describe this approach, we consider the problem of estimating the quality of a given patch. Assume that the encounter of food items within the patch is a Poisson process with parameter λ, as in equation (1) above. To keep the argument simple, we assume here that λ is constant within a patch, but not from patch to patch. Depletion of the patch is considered in the Appendix.

The forager does not know the value of λ when it enters the patch. However, it has a prior expectation of patch quality obtained from previous experience. Let $f(\lambda)$ denote this prior probability distribution (i.e., $f(\lambda)d\lambda$ is the probability that the patch quality parameter lies between λ and $\lambda + d\lambda$). To be explicit, we shall assume that $f(\lambda)$ is a gamma distribution, although other distributions could equally be chosen (see Iwasa et al. 1981; Mangel and Clark 1983):

$$f(\lambda) = \gamma(\lambda;\nu,\alpha) = \frac{\alpha^\nu}{\Gamma(\nu)}\lambda^{\nu-1}e^{-\alpha\lambda} \qquad (\lambda \geq 0). \tag{7}$$

The mean and variance of this distribution are:

$$E\{\lambda\} = \nu/\alpha \quad \text{and} \quad \mathrm{Var}\{\lambda\} = \nu/\alpha^2 \tag{8}$$

and hence the coefficient of variation is $\mathrm{CV}\{\lambda\} = 1/\sqrt{\nu}$. Here ν is a dimensionless parameter and α has units of time.

Assume that the forager enters the patch, searches for a time t and locates n

food items. It is easily shown by an application of Bayes' theorem (see DeGroot 1970) that the updated probability distribution for λ is again a gamma distribution, but with updated parameters ν' and α' given by

$$\nu' = \nu + n, \qquad \alpha' = \alpha + t. \tag{9}$$

The expected encounter rate $E\{\lambda\}$ thus changes to

$$E\{\lambda | n,t\} = \frac{\nu + n}{\alpha + t} \tag{10}$$

while the coefficient of variation declines to $1/\sqrt{\nu + n}$.

The analogue of equation (10) when patch depletion is included is discussed in the Appendix; see also Iwasa et al. (1981), who obtain updating formulas under depletion for the case in which the distribution of the number of prey per patch is a negative binomial, binomial, or Poisson distribution.

How can the forager utilize the updated estimates? The optimal foraging literature has mostly addressed the question of the optimal length of time to remain in a given patch (the "residence time" or "giving-up time"; see McNair [1982]). Under conditions of uncertainty this formulation of the optimal foraging problem is overly restrictive. A more appropriate question would be, How should the forager allocate his limited resources (i.e., time) between the available patches, or feeding areas, so as to maximize some specified objective or combination of objectives?

Consider, for example, the case of two patches, with prior distributions $f_i(\lambda_i)$, $i = 1,2$. Would the smart forager choose one patch, feed in it for a certain time (depending on its feeding success there), and then switch to spend the remaining time (possibly zero) in the second patch? Or would the forager sample both patches before deciding.

Not surprisingly, it turns out in many cases to be extremely important to sample both patches before deciding (Mangel and Clark 1983). The greater the initial uncertainty (as measured by the coefficient of variation of the prior distribution f_i), the more important it is to sample the patch. Once sampling has reduced the coefficient of variation, the more profitable patch should then be exploited. Of course, if patch quality varies randomly over time, sampling presumably should be continued, although this case has not been analyzed mathematically.

It seems likely that many foragers do recognize the importance of sampling, although most experimental results reported in the literature seem to exclude the data from any initial sampling phase on the part of foragers. Pyke et al. (1977), however, report experiments in which birds did not spend all their foraging time on patches of greatest abundance, but regularly probed other patches.

Similar considerations apply to other aspects of the foraging problem, such as estimating the abundance of patches in the environment. For simplicity, we now adopt a deterministic model for feeding within a patch:

$$B_t = B_0(1 - e^{-ft}) \tag{11}$$

where B_t denotes the total amount of food obtained from the patch in t units of time, B_0 is the initial amount of food, and f is the proportional feeding rate. Here

B_0 and f are assumed known, and are the same for all patches. The instantaneous rate of consumption is then

$$\frac{dB_t}{dt} = B_0 f e^{-ft}. \tag{12}$$

Assume that the search for patches is a Poisson process with parameter λ. The expected time to find a patch, given λ, is then

$$T_\lambda = E\{t|\lambda\} = 1/\lambda. \tag{13}$$

As before, let us assume that the prior probability for λ is a gamma density, as given by equation (7). Then the prior expectation of T_λ, as found in the Appendix, is

$$\langle T_\lambda \rangle = \frac{\alpha}{\nu - 1}. \tag{14}$$

The prior expected feeding rate for the environment is given by

$$\frac{B_0(1 - e^{-ft})}{t + \langle T_\lambda \rangle} \tag{15}$$

where t is the amount of time spent in each patch (the residence time RT). If the forager does not update information on its environment, it will leave a patch at the time t which equates the expressions (12) and (15). Upon simplification, the equation for t becomes

$$f[t + \langle T_\lambda \rangle] = e^{ft} - 1. \tag{16}$$

It is easily seen that this equation determines a unique positive residence time t, which is an increasing function of $\langle T_\lambda \rangle$, i.e., the longer the expected search time, the longer the forager remains in each patch. A similar result was obtained by Charnov (1976).

Suppose, however, that the forager uses the information obtained from searching for patches to update his estimate of patch abundance. If t_1 denotes the time spent finding the first patch, then the updated value of λ is

$$\lambda' = E\{\lambda|1,t_1\} = \frac{\nu + 1}{\alpha + t_1}. \tag{17}$$

Thus $\langle T_{\lambda'} \rangle = (\alpha + t_1)/\nu$, and this should be used in equation (16) in place of $\langle T_\lambda \rangle$.

Let t and t' denote the RT's as calculated from equation (16) without updating, and with updating after the first encounter, respectively. Table 1 shows the percentage change in the RT, $(t' - t)/t$, resulting from this updating, for a variety of parameter values. Whenever the observed encounter time t_1 is less than the anticipated value $\langle T_\lambda \rangle$, the updating forager decreases his RT, and vice versa. The percentage change in RT depends on how unexpected t_1 is, and also on the prior distribution parameters ν, α.

The increase in feeding rate resulting from this initial update varies from 0% to 7% for the range of values covered in the table. Further updates, which would be made as subsequent patches were located, would increase the overall feeding rate.

TABLE 1

PROPORTIONAL CHANGE IN RESIDENCE TIMES, $(t' - t)/t$, CAUSED BY UPDATING THE ESTIMATED ENVIRONMENTAL QUALITY AFTER THE FIRST PATCH IS ENCOUNTERED
(Proportional feeding rate $f = 0.1$ per unit time.)

$t_1/\langle T_\lambda \rangle$	$\nu = 1.1$, $\alpha = .5$, $\langle T_\lambda \rangle = 5$	$\nu = 1.1$, $\alpha = .05$, $\langle T_\lambda \rangle = .5$	$\nu = 5$, $\alpha = 1$, $\langle T_\lambda \rangle = .25$
.2	− .44	− .46	− .08
.4	− .29	− .31	− .06
.6	− .18	− .19	− .04
.8	− .08	− .09	− .02
1.0	0	0	0
1.2	+ .07	+ .08	+ .02
1.4	+ .14	+ .16	+ .04
1.6	+ .20	+ .23	+ .06
1.8	+ .26	+ .29	+ .07
2.0	+ .31	+ .36	+ .09

THE INFORMATIONAL ADVANTAGES OF FLOCKING

We have shown (1) that flocking may be the evolutionarily stable strategy in terms of the individual's feeding rate, and also (2) that informational updating is important for the individual.

Flocking may confer an additional evolutionary advantage inasmuch as flocking results in more rapid updating of estimates of patch and environmental quality than does individual searching. On the other hand, a flock depletes the local environment more rapidly than does an individual animal. The following simple model captures these two countereffects.

Suppose there are two types of patches, "good," containing a useful quantity B of food, and "bad," containing a negligible amount of food. The good patches occur with probability p and the bad with probability $1 - p$. Let τ denote the average time required for an animal to decide whether a particular patch is good or bad, and let T be the average time required to eat a good patch. The average feeding rate for the individual animal is then

$$\frac{pB}{pT + \tau} \tag{18}$$

(time spent looking for patches is ignored).

A flock of n animals will both search and eat at n times the individual rate. Hence the flock's decision and eating times per patch are τ/n and T/n, and the amount eaten per individual per patch is B/n. Consequently the average feeding rate per individual is unchanged from equation (18). Thus (as in our first model) there appears to be no informational advantage to flocking—at least as far as the average feeding rate is concerned. In fact, second-order effects would imply a disadvantage to flocking from this point of view.

Suppose now that $p \ll 1$. Then the individual encounters good patches infrequently, although it finds plenty to eat (B) when it finds a patch. As a member of a

flock of n animals, searching independently but communicating their discoveries, the individual encounters patches n times as often, but eats only B/n per discovery. Thus the variation in the feeding rate is reduced by flocking. This obviously may be important to the individual, especially when patches are rare.

We have thus arrived at our second advantage of foraging by flocks in a patchy environment, reduction in the variation of individuals' feeding schedules brought about by the increased use of information (see Caraco 1981).

Finally, the rate at which information is generated by cooperative foraging is clearly greater than for individual foraging. To make this more precise, consider a given feeding area, in which search for food items is a Poisson process with parameter λ, where λ has a gamma prior distribution, as above. The probability that a flock of n foragers locates k food items in time t is then (Mangel and Clark 1983; also see Feller [1957] who discusses this negative binomial distribution)

$$p(k,n,t) = \frac{(nt)^k}{k!} \frac{\alpha^\nu}{(\alpha + nt)^{\nu+k}} \frac{\Gamma(\nu + k)}{\Gamma(\nu)}. \tag{19}$$

After k food items have been encountered, the coefficient of variation of the updated (posterior) distribution becomes $1/\sqrt{\nu + k}$. We define the prior expected coefficient of variation $E\{\mathrm{CV}|n,t\}$ by

$$E\{\mathrm{CV}|n,t\} = \sum_{k=0}^{\infty} \frac{1}{\sqrt{\nu + k}} p(k,n,t). \tag{20}$$

This expression gives an a priori estimate of the accuracy of the information obtained by n foragers in time t. Note that $E\{\mathrm{CV}|n,t\} = E\{\mathrm{CV}|1,nt\}$, i.e., n birds searching for time t achieve the same accuracy as one bird in time nt, as we assumed in the simple arguments presented above.

A plot of the prior expected coefficient of variation versus flock size n is shown in figure 3. As expected, the coefficient of variation is a decreasing function of n.

SUMMARY

In this paper we have investigated the effects of uncertainty and information on the foraging strategies of animals. We first argue that group foraging, or flocking, can increase individual feeding rates as a result of the sharing of information among group members. Also, members of a foraging group may experience reduced variation in their feeding rates. Both of these advantages increase with the scarcity and patchiness of the food supply. On the other hand, our analysis also suggests that the informational advantages of flocking may tend to be dissipated by "overflocking," which is in fact an evolutionarily stable strategy. (Possible defense strategies against overflocking, such as territoriality or aggression, have not been considered here.

We also discuss the strategy of individual foragers. We suggest that the traditional analysis in terms of marginal values, residence times, and giving-up times, may be misleading. It turns out to be important to "probe," especially in patches where food abundance is highly variable. The marginal value theorem fails completely to reflect this aspect of uncertainty and information.

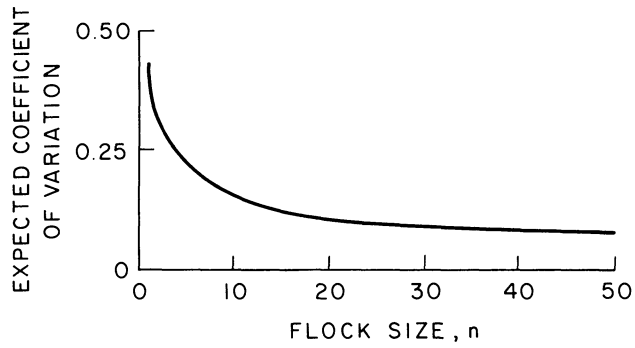

FIG. 3.—Expected coefficient of variation in estimated patch size λ_i as a function of flock size. (Parameter values: $\nu = 1$, $\alpha = 0.1$, $T = 1$.)

ACKNOWLEDGMENTS

We thank A. Stewart-Oaten, University of California, Santa Barbara, for bringing his work to our attention. We are indebted to S. A. Levin of Cornell University, and an anonymous reviewer, for a number of corrections and improvements in the paper. The work of M. M. was partially supported by NSF Grant MCS-21659, and that of C. W. C. by NSERC Grant A-3990.

MATHEMATICAL APPENDIX

First we discuss certain matters of a probabilistic nature; most of the calculations are omitted, being available in elementary texts (e.g., Feller 1957, 1971).

The Poisson process, defined by equation (1), is our basic model of searching for patches, or for food items within a patch; the parameter λ is the encounter rate, and (for a given searcher) is thus proportional to the abundance of patches (or of food items in a patch). If n foragers search independently (possibly sharing information on their discoveries), the encounter rate becomes $n\lambda$.

The probability that one searcher makes exactly k encounters in a time interval of length t, given λ, is

$$p_k(\lambda,t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0,1,2, \ldots . \tag{21}$$

The expected time to encounter one patch is easily seen to be

$$T = 1/\lambda. \tag{22}$$

Since λ is assumed to be constant the Poisson process applies only to the case in which items encountered are not removed. The Poisson process is thus unsuitable for the study of resource depletion, in which case equation (1) should be replaced by (Mangel and Clark 1983):

$$\begin{aligned}&\text{Prob (one encounter in time } dt\text{, given that}\\ &\quad i \text{ items have already been encountered)}\\ &\qquad = (\lambda - i\delta)dt.\end{aligned} \tag{23}$$

Here λ denotes the initial encounter rate, and $\lambda/\delta = B_0$ is the total number of items present initially. We assume that the parameter δ is known, but λ may be uncertain. Given λ, the

probability that one searcher encounters and removes k items in time t is

$$F_k(\lambda,t) = \binom{\lambda/\delta}{k}(1 - e^{-\delta t})^k e^{-(\lambda - k\delta)t} \qquad (k \leq \lambda/\delta). \tag{24}$$

The expected number of items encountered in time t is then

$$\overline{F}(\lambda,t) = \frac{\lambda}{\delta}(1 - e^{-\delta t}) \tag{25}$$

which is the same as equation (11).

Next we consider the updating formulas. The prior distribution for the search parameter λ is taken to be the gamma distribution

$$f(\lambda) = \gamma(\lambda;\nu,\alpha) = \frac{\alpha^\nu}{\Gamma(\nu)}\lambda^{\nu-1}e^{-\alpha\lambda} \tag{26}$$

where $\Gamma(\nu) = \int_0^\infty e^{-s}s^{\nu-1}\,ds$ is the gamma function. If $f(\lambda|k,t)$ denotes the updated probability distribution for λ after k items have been encountered in time t, we have by Bayes' theorem

$$f(\lambda|k,t) = \frac{\mathrm{pr}(k|\lambda,t)f(\lambda)}{\displaystyle\int_0^\infty \mathrm{pr}(k|\mu,t)f(\mu)d\mu} \tag{27}$$

where $\mathrm{pr}(k|\lambda,t)$ is the probability of encountering k items in time t, given λ.

In the case of Poisson search (without removals) we have $\mathrm{pr}(k|\lambda,t) = p_k(\lambda,t)$ as in equation (21). Carrying out the integration in equation (27) and simplifying, we obtain

$$f(\lambda|k,t) = \gamma(\lambda;\ \nu + k,\ \alpha + t) \tag{28}$$

which is the updating formula used in the text.

If the removal of encountered items is to be taken into account, $\mathrm{pr}(k|\lambda,t)$ must be equated to the expression $F_k(\lambda,t)$ in equation (24). The resulting formula for the updated distribution $f(\lambda|k,t)$ becomes somewhat messy, but its values are easily calculated on a small computer (Mangel and Clark 1983).

The prior expected value of any random variable $r(\lambda)$ is obtained by integrating $r(\lambda)$ against the prior distribution $f(\lambda)$. For example, we have

$$\langle T_\lambda \rangle = \int_0^\infty \frac{1}{\lambda}\gamma(\lambda;\nu,\alpha)d\lambda = \frac{\alpha}{\nu - 1} \qquad (\nu > 1). \tag{29}$$

(A quirk of the gamma distribution is that the prior expected time to encounter one item becomes infinite for $\nu \leq 1$, i.e., if the coefficient of variation is ≥ 1, regardless of the mean.)

Finally, we describe the analysis used to construct the feeding rate versus flock-size curves shown in figure 1. The setup is the following. The flock discovers food patches according to a Poisson process with parameter $n\lambda$ that depends upon flock size n. A patch of food is assumed to contain m meals (which is, for the time being, assumed to be deterministic). The time required for the flock to consume the patch is $\tau(n)$, after which searching resumes. We assume that $\tau(n)$ has the form

$$\tau(n) = \begin{cases} \tau_1 & \text{if } n = 1 \\ \tau_1(1 + \epsilon\sqrt{n}) & \text{if } 2 \leq n \leq m \\ \tau_1 \dfrac{m}{n}(1 + \epsilon\sqrt{n}) & \text{if } n > m. \end{cases} \tag{30}$$

The term $\epsilon\sqrt{n}$ captures the "crowding" effects of large flocks in an ad hoc fashion. Our assumption is that the n birds in a flock search independently and that as long as flock size is less than the number of meals per patch, the birds spend τ_1 units of time eating. As n

increases past m, the eating time decreases. An individual bird in the flock will receive one meal per patch if $n \leq m$ and m/n of a meal per patch if $n > m$.

The average feeding rate can be determined in a simple fashion by applying renewal theory (Mangel 1982). Consider a time period T, much greater than $1/\lambda$ or τ_1. The process search-eat-search is a renewal process (Feller 1957). Let $\bar{P}(T)$ be the expected number of patches found in time T. Applying the renewal theorem gives

$$\bar{P}(T) \sim \frac{n\lambda T}{n\lambda\tau(n) + 1} - \frac{n\lambda\tau(n)[n\lambda\tau(n) + 2]}{2[n\lambda\tau(n) + 1]^2} + 0\left(\frac{1}{T}\right). \tag{31}$$

We view the quotient $\bar{P}(T)/T$ as the long-run rate of finding patches of food. Let $\theta(n,m)$ be 1 if $n < m$ and m/n otherwise. Then the long-run feeding rate $r(n)$ of a single bird in a flock of size n is

$$r(n) = \theta(n,m)\frac{n\lambda}{n\lambda\tau(n) + 1} - \frac{1}{2T}\frac{n\lambda\tau(n)[n\lambda\tau(n) + 2]}{[n\lambda\tau(n) + 1]^2} + 0\left(\frac{1}{T^2}\right). \tag{32}$$

The first term has the following interpretation: dividing top and bottom by $n\lambda$ gives $1/[\tau(n) + 1/n\lambda]$. The denominator in this expression is the average total time spent in one search-eat cycle. In case $\epsilon = 0$, this term reduces to the expression used in equation (3).

If T is large, only the first term in equation (32) will contribute significantly to the result. The curve shown in figure 1 used both terms, for a moderate value of T.

In nature, the actual number of meals per patch is certainly a random variable. This can easily be included in the model by averaging over some assumed probability distribution for m, but this modification results in no qualitative change to the model.

LITERATURE CITED

Caraco, T. 1981. Risk sensitivity and foraging groups. Ecology 62:527–531.

Caraco, T., and L. L. Wolf. 1975. Ecological determinants of group sizes of foraging lions. Am. Nat. 109:343–352.

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

Clark, C. W. 1983. The effect of fishermen's quotas on expected catch rates. J. Mar. Res. Econ. (in press).

Cody, M. L. 1971. Finch flocks in the Mohave desert. Theor. Popul. Biol. 2:142–158.

Cowie, R. J., and J. R. Krebs. 1979. Optimal foraging in patchy environments. Pages 183–205 in B. D. Turner and L. R. Taylor, eds. Population dynamics, 20th Symposium of the British Ecological Society. Blackwell Scientific, Oxford.

DeGroot, M. 1970. Optimal statistical decisions. McGraw-Hill, New York.

Feller, W. 1957 and 1971. An introduction to probability theory and its applications. Vols. I, II. Wiley, New York.

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

Hardin, G. 1968. The tragedy of the commons. Science 162:1243–1248.

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

Krebs, J. R. 1978. Optimal foraging: decision rules for predators. Pages 23–63 *in* J. R. Krebs and N. B. Davies, eds. Behavioral ecology: an evolutionary approach. Sinauer, Sunderland, Mass.

Krebs, J. R., and N. B. Davies, eds. 1978. Behavioral ecology: an evolutionary approach. Sinauer, Sunderland, Mass.

McNair, J. N. 1982. Optimal giving-up times and the marginal value theorem. Am. Nat. 119:511–529.

Mangel, M. 1982. Search, effort, and catch rate in fisheries. Eur. J. Oper. Res. 11:361–366.

Mangel, M., and C. W. Clark. 1983. Uncertainty, search, and information in fisheries. J. Cons. Int. Explor. Mer 41:93–103.

Maynard Smith, J. 1974. The theory of games and the evolution of animal conflicts. J. Theor. Biol. 47:209–221.

Nash, J. F. 1951. Non-cooperative games. Ann. Math. 54:286–295.

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

Pulliam, H. R., G. H. Pyke, and T. Caraco. 1982. The scanning behavior of juncos: a game-theoretical approach. J. Theor. Biol. 85:89–104.

Pyke, G. H., H. R. Pulliam, and E. L. Charnov. 1977. Optimal foraging: a selective review of theory and tests. Q. Rev. Biol. 52:137–154.