The Evolution of Bayesian Updating

Samir Okasha*†

An evolutionary basis for Bayesian rationality is suggested, by considering how natural selection would operate on an organism's 'policy' for choosing an action depending on an environmental signal. It is shown that the evolutionarily optimal policy, as judged by the criterion of maximal expected reproductive output, is the policy that, for each signal, picks an action that maximizes conditional expected output given that signal. This suggests a possible route by which Bayes-rational creatures might have evolved.

1. Introduction. Bayesian updating, also known as 'conditionalization', is a rule specifying how a prior probability distribution should be updated to a posterior distribution in the light of new information. The rule is often interpreted epistemically: it specifies how an agent should change its epistemic state over time in response to new evidence, where an 'epistemic state' is represented by a probability distribution over some specified set of alternatives.

There has been considerable philosophical discussion of the normative status of Bayesian updating. Does rationality require an agent to update using the Bayesian rule? If so, can the rule be justified from more primitive rationality requirements? Is it the unique rule that can be so justified? Are there situations in which Bayesian updating should not be used? Works addressing these questions include Brown (1976), van Fraassen (1989), Christensen (1991), Maher (1992), Lewis (1999), Greaves and Wallace (2006), and Leitgeb and Pettigrew (2010).

Here I focus on a different although related issue. Instead of asking whether Bayesian updating is a requirement of rationality, I ask whether

*To contact the author, please write to: Department of Philosophy, University of Bristol; e-mail: Samir.Okasha@bristol.ac.uk.

†Thanks to John McNamara, Cedric Paternotte, Jonathan Grose, Chris Clarke, Richard Pettigrew, James Ladyman, and Nick Shea for discussion. This work was supported by the European Research Council Seventh Framework Program (FP7/2007-2013), ERC grant 295449.

Philosophy of Science, 80 (December 2013) pp. 745–757. 0031-8248/2013/8005-0044$10.00
Copyright 2013 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/674058 Published online by Cambridge University Press

it is a requirement of evolutionary optimality. Changing one's beliefs in response to evidence is an aspect of cognition, and it seems likely that at least some aspects of cognition, animal and human, have been shaped by natural selection. This prompts the question of how selection would have operated on updating rules. An organism's updating rule affects its epistemic state, which in turn affects its behavior, so different updating rules will lead to different behaviors and, thus, have different consequences for survival and reproduction. Can we show that organisms using Bayesian updating would have enjoyed a selective advantage over those updating in some other way?

This question is an aspect of the broader issue of whether natural selection will tend to produce rational behavior. A number of authors have examined whether rational-choice norms such as transitivity of preference, maximization of expected utility, and consistency of time preference can be derived from an underlying evolutionary model (Skyrms 1995; Robson 1996; Cooper 2003; Houston, McNamara, and Steer 2007). However, the question of whether Bayesian updating can be derived from evolutionary principles has never been considered, so far as I know. Here I sketch a tentative answer to that question.

2. Bayesianism in Evolutionary Biology. Some readers may be surprised by our discussing Bayesian principles in an evolutionary context, on the grounds that nonhuman animals lack the cognitive sophistication required to have probabilistic beliefs about the world, less still to update them. But in fact, behavioral ecologists make widespread use of Bayesian ideas to think about animal behavior, especially in foraging theory (e.g., McNamara, Green, and Olsson 2006).

A typical Bayesian approach in foraging theory is to assume that animals have some prior information about an environmental parameter, represented by a probability distribution. For example, the animal may 'know' that food patches are of two types, good and bad, whose relative frequencies are 1/3 and 2/3 respectively. This information may be either genetically encoded or acquired by the animal through its experience. As the animal begins to forage, it updates its belief about the type of patch it is in. If the probability that it is in a bad patch becomes sufficiently high, the animal may choose to move on. Many optimal foraging models aim to derive a precise prediction about when an animal should move from one patch to another.

In models of this sort, talk of an animal's 'knowledge' is intended behavioristically. Saying that the animal 'knows' that 1/3 of food patches are good means that the animal behaves as if it knows this, as manifested in its behavior. If the environment can be in a number of different states, and the appropriate action depends on the state, then to behave adaptively an

https://doi.org/10.1086/674058 Published online by Cambridge University Press

animal will need to be sensitive to the states' probabilities of occurrence. If the animal chooses between actions according to maximization of expected reproductive output—one standard definition of adaptive behavior—then it is behaving as if it knows the true probability distribution on the states, knows the payoffs, and is able to calculate expectations.

Understanding 'knowledge' in this behavioristic way may seem quite different from how the notion is understood in epistemology. But in fact many Bayesian decision theorists, notably Savage (1954), have argued for a behavioristic interpretation of probabilistic beliefs, according to which an agent's subjective probabilities (and utilities) are derived from their choices between uncertain prospects, which are in principle observable; this interpretation is orthodox in contemporary economics, if not in contemporary epistemology. So in fact, evolutionary biologists' use of Bayesian concepts to model animal behavior is not such a radical departure.

3. A Simple Example. To fix ideas, consider a simple example. An organism is foraging for food in a predator-strewn area. Predators are of two types, snakes and leopards. If no predator is present, the best thing to do is stay put and forage. If a snake is present, the best thing to do is to climb a tree. If a leopard is present, the best thing to do is flee. Both climbing and fleeing are costly in terms of time and energy, to different extents. Payoffs for the three actions, as a function of the state of the world, are shown in table 1; these payoffs are measured in increments of biological fitness (i.e., number of offspring).

What action should the animal choose to maximize its expected payoff? That depends on the probabilities of the three states of the world. Suppose that the probabilities of the three states are

$$\mathrm{P}(\text{no predator}) = \frac{1}{2}, \quad \mathrm{P}(\text{snake}) = \frac{1}{3}, \quad \mathrm{P}(\text{leopard}) = \frac{1}{6}.$$

(These probabilities can be thought of as the relative frequencies with which each state occurs in the relevant local ecology.) Then, the expected payoff from each action is

TABLE 1. Payoffs for Three Alternative Actions

	No Predator	Snake	Leopard
Stay put	10	0	0
Climb	5	4	0
Flee	6	1	2

https://doi.org/10.1086/674058 Published online by Cambridge University Press

$$\text{V(stay put)} = 10\left(\frac{1}{2}\right) + 0\left(\frac{1}{3}\right) + 0\left(\frac{1}{6}\right) = 5,$$

$$\text{V(climb)} = 5\left(\frac{1}{2}\right) + 4\left(\frac{1}{3}\right) + 0\left(\frac{1}{6}\right) = 3.83,$$

$$\text{V(flee)} = 6\left(\frac{1}{2}\right) + 1\left(\frac{1}{3}\right) + 2\left(\frac{1}{6}\right) = 3.67.$$

So the evolutionarily optimal action is to stay put. Organisms choosing to stay put will on average leave more offspring than those choosing either of the other two actions. If the organisms' choice behavior has been optimized by natural selection, or if it has learned which action is optimal, then it will stay put.

Now suppose that before choosing an action, the organism receives a signal that indicates whether a predator is present. So the signal has two values: 'safe' and 'unsafe'. The signal is perfectly reliable (i.e., indicates 'unsafe' if and only if a predator is present) but cannot discriminate between leopards and snakes. The organism's choice of action may depend on which signal is received. So the organism needs to have a 'policy', that is, a specification of which action to take for each value of the signal. There are 9 (3^2) possible policies, in this example. Let us consider three policies in particular.

Suppose first that the organism ignores the signal and always chooses the action optimal for the situation in which no signal is received (i.e., stay put). So the organism's policy is 'if safe, stay put; if unsafe, stay put'. Let us denote this policy 'Ignore'. The policy may seem unpromising, intuitively, but it could reflect the organism's cognitive limitations. An organism incapable of attending to the signal, or lacking behavioral plasticity, might use a policy like Ignore.

Second, suppose that the organism does attend to the signal and adopts a maximin strategy; that is, it chooses the action that maximizes the minimum payoff it will receive, in the light of the signal. So if the safe signal is received, it chooses to stay put, obviously. If the unsafe signal is received, it chooses to flee, as this guarantees it a payoff of at least 1. So its policy is 'if safe, stay put; if unsafe, flee'. Let us call this policy 'Maximin'. The policy reflects a high degree of risk aversion. By choosing to flee rather than climb when the unsafe signal is received, the organism forgoes a possible payoff of 4 in order to definitely avoid a payoff of 0.

Third, suppose that the organism behaves like a Bayesian. On receipt of a signal, the organism chooses an action that maximizes its conditional expected payoff, given the signal.[1] So if the safe signal is received, the organism chooses to stay put—as the conditional expected payoffs are then

1. Note that this action need not be unique, although it is in this example; see sec. 4.

https://doi.org/10.1086/674058 Published online by Cambridge University Press

10, 5, and 6 for stay put, climb, and flee respectively. (This is because the conditional probability of the no-predator state, given the safe signal, is 1.) What if the unsafe signal is received? The conditional probabilities of the no-predator, snake, and predator states, given an unsafe signal, are 0, 2/3, and 1/3 respectively. The conditional expected payoffs for staying put, climbing, and fleeing are then 0, 8/3, and 4/3 respectively, so the organism will choose to climb. Its policy is therefore 'if safe, stay put; if unsafe, climb'. Let us call this policy 'Bayes'.

Which of our three policies—Ignore, Maximin, and Bayes—is the best, from an evolutionary point of view? To answer this question, we need to compute the expected payoffs accruing to an organism that uses each of the policies. Consider first Ignore. Since an organism using Ignore chooses to stay put whether or not a signal is received, its expected payoff is

$$\mathrm{V}[\text{Ignore}] = 10\left(\frac{1}{2}\right) + 0\left(\frac{1}{3}\right) + 0\left(\frac{1}{6}\right) = 5.$$

What about Maximin? An organism using Maximin stays put if a safe signal is received but flees otherwise. With probability 1/2, no predator is present, so the safe signal is sent, so the organism stays put and earns a payoff of 10. With probability 1/3, a snake is present, so the unsafe signal is sent, so the organism flees and earns a payoff of 1. With probability 1/6, a leopard is present, so the unsafe signal is sent, so the organism flees and earns a payoff of 2. Its expected payoff is, therefore,

$$\mathrm{V}[\text{Maximin}] = 10\left(\frac{1}{2}\right) + 1\left(\frac{1}{3}\right) + 2\left(\frac{1}{6}\right) = 5.67.$$

By a similar logic, the payoff to an organism using the Bayes policy is

$$\mathrm{V}[\text{Bayes}] = 10\left(\frac{1}{2}\right) + 4\left(\frac{1}{3}\right) + 0\left(\frac{1}{6}\right) = 6.33.$$

Therefore, the Bayes policy yields the highest expected payoff. If all three policies are found in a population, natural selection will favor the Bayes policy over the other two. This suggests, obviously in a preliminary way, that evolution might produce organisms that behave as if they were rational Bayesian agents, whose subjective probabilities over the states of nature match the objective frequencies, and who incorporate new information by conditionalization. A stronger argument for this conclusion

https://doi.org/10.1086/674058 Published online by Cambridge University Press

would need to show that no conceivable policy does better than the Bayes policy; see section 4.

4. Cannot Do Better than Bayes. The framework we adopt directly generalizes the previous example. There is a finite set S of states of nature; $S = \{\theta_1, \ldots, \theta_k\}$. There is a finite set U of actions that the organism may perform. The payoff from an action depends on the state of nature. The payoff from action $u \in U$ if the state of nature turns out to be θ_i is $V_i(u)$. As before, payoffs are measured in increments of biological fitness.

There is a finite set E of possible signals, or items of evidence, that the organism may receive; $E = \{E_1, \ldots, E_n\}$. In the example above we assumed that the signal set partitioned the states of nature, that is, each state of nature was compatible with exactly one signal, but no such assumption is made here. There is a joint probability distribution p on $S \times E$, reflecting the frequency with which each (state, signal) pair occurs. The marginal probability of state θ_i is $p(\theta_i)$, and that of signal E_j is $p(E_j)$. We assume that $p(\theta_i) > 0$ for all i and $p(E_j) > 0$ for all j; that is, each state occurs with positive probability, and similarly for each signal.

An organism's 'policy' specifies an action $u \in U$ for every possible signal $E_j \in E$. Thus, a policy is simply a function from E to U. The set of all policies is denoted X. For any policy $x \in X$, we let $x(E_j) \in U$ be the action specified by that policy when signal E_j is received; for convenience, we write $x(E_j)$ as x_j.

Consider an arbitrary policy $x \in X$. What is an organism's expected payoff from using policy x? Suppose first that the true state of nature is θ_i. Then, the expected payoff from policy x equals

$$\sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(x_j). \tag{1}$$

The justification for (1) is clear: $p(E_j/\theta_i)$ is the probability that signal E_j is received given that the state of nature is θ_i; $V_i(x_j)$ is the payoff from choosing action x_j—the action specified by policy x when the signal received is E_j—in state of nature θ_i. Expression (1) is thus the expected payoff to an organism using policy x when the true state is θ_i. Taking the expectation of (1) across states of nature thus gives us the overall **expected payoff to policy x**:

$$\sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(x_j). \tag{2}$$

Now recall the Bayes policy discussed above. In our simple example above, there was a unique Bayes policy since for each signal there was a

https://doi.org/10.1086/674058 Published online by Cambridge University Press

unique action that maximized conditional expected payoff given the signal. But in the general case this need not be so: for some signals, there may be more than one action satisfying this maximization constraint. So we need to consider the class of Bayes policies $B \subset X$, where each $b \in B$ is a policy that, for each signal $E_j \in E$, chooses an action $b_j \in U$ that maximizes conditional expected payoff given E_j.

We wish to show that a policy is evolutionarily optimal if and only if it is a Bayes policy; as before, 'evolutionarily optimal' means 'earns an expected payoff greater than or equal to every other policy'. To show this, consider any Bayes policy $b \in B$. Substituting b for x in expression (2) thus gives us the **expected payoff to a Bayes policy**:

$$\sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j). \tag{3}$$

Our first task is to show that expression (3) is greater than or equal to (2); that is, policy b has expected payoff greater than or equal to policy x.

Since b is a Bayes policy, we know that it satisfies the following condition: for every signal E_j,

$$\sum_{\theta_i} p\left(\frac{\theta_i}{E_j}\right) V_i(b_j) \geq \sum_{\theta_i} p\left(\frac{\theta_i}{E_j}\right) \times V_i(u), \tag{4}$$

for all actions $u \in U$. Applying inequality (4) to the particular action $x_j \in U$ gives

$$\sum_{\theta_i} p\left(\frac{\theta_i}{E_j}\right) \times V_i(b_j) \geq \sum_{\theta_i} p\left(\frac{\theta_i}{E_j}\right) \times V_i(x_j). \tag{5}$$

Applying Bayes's theorem and dividing across by $p(E_j)$,

$$\sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j) \geq \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(x_j). \tag{6}$$

Summing over all the signals,

$$\sum_{E_j} \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j) \geq \sum_{E_j} \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(x_j). \tag{7}$$

Reversing the order of summation,

$$\sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j) \geq \sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(x_j). \tag{8}$$

https://doi.org/10.1086/674058 Published online by Cambridge University Press

But the left-hand side of (8) is the expected payoff to policy b (eq. [3]), while the right-hand side is the expected payoff to policy x (eq. [2]). So policy b is evolutionary optimal. Since nothing has been assumed about policy b except that it is a Bayes policy, we can conclude that every Bayes policy is evolutionarily optimal.

To show the converse, consider another policy y, which is not a Bayes policy. Thus, for some signal E_c,

$$\sum_{\theta_i} p\left(\frac{\theta_i}{E_c}\right) \times V_i(y_c) < \sum_{\theta_i} p\left(\frac{\theta_i}{E_c}\right) \times V_i(b_c). \tag{9}$$

Applying Bayes's theorem,

$$\sum_{\theta_i} p(\theta_i) \times \left(\frac{E_c}{\theta_i}\right) \times V_i(y_c) < \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_c}{\theta_i}\right) \times V_i(b_c). \tag{10}$$

Since b is a Bayes policy, we know from (6) that for all signals E_j:

$$\sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(y_j) \leq \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j). \tag{11}$$

From (10) and (11), we can deduce that, summing over all the signals,

$$\sum_{E_j} \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(y_j) < \sum_{E_j} \sum_{\theta_i} p(\theta_i) \times p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j). \tag{12}$$

Reversing the order of summation,

$$\sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(y_j) < \sum_{\theta_i} p(\theta_i) \sum_{E_j} p\left(\frac{E_j}{\theta_i}\right) \times V_i(b_j). \tag{13}$$

But the left-hand side of (13) is the expected payoff to policy y, while the right-hand side is the expected payoff to the Bayes policy b. So any policy that is not a Bayes policy earns a strictly lower expected payoff than a Bayes policy, so it cannot be evolutionarily optimal. Therefore, an evolutionarily optimal policy must be a Bayes policy.

Taken together, (8) and (13) tell us that a policy is evolutionarily optimal if and only if it is a Bayes policy. So organisms that implement Bayes policies will leave more offspring, on average, than ones that do not. Over time, we would expect the non-Bayes policies to be eliminated by natural selection.

5. Discussion. What exactly does the foregoing argument show? The argument seems to supply an evolutionary basis for Bayesian updating, by

https://doi.org/10.1086/674058 Published online by Cambridge University Press

showing that it follows from the standard Darwinian assumption that animal behavior is well adapted and, hence, designed to maximize the animal's expected reproductive output. However, some care is needed here. Typical discussions of Bayesian updating ask how an agent does or should update its epistemic state on receipt of information, where an epistemic state is represented by a prior probability distribution over some appropriate algebra. But our argument above makes no mention of 'epistemic states' or ways of updating them, and it makes no assumption that the organisms in question are capable of being in such states. We have assumed only that the organisms are behaviorally plastic and capable of modifying their behavior in response to an environmental signal.

To get from this behaviorist starting point to something more epistemic, we need to consider how an organism might implement an evolutionarily optimal policy, as characterized above. One possibility is this: the organism perates with an internal 'belief-like' representation of the environment and an internal 'desire-like' representation of the outcomes (or payoffs), and it uses a 'choice rule' to determine which action to pick, depending on which belief and desire states it is in. On receipt of a signal, the organism uses an 'update rule' to go from one belief state to another. Given a generic psychological makeup of this sort, which is presumably applicable to at least some higher animals as well as humans, an evolutionarily optimal policy is straightforward to implement.

To see this, suppose we take a 'belief state' to be a subjective probability distribution over the states of nature and signals and a 'desire state' to be a real-valued utility function on the outcomes (or action-state pairs). In particular, suppose that the organism's initial belief state, before receiving a signal, is equal to the true prior distribution p, and its utility function is equal to the true payoff function V (or an affine transformation thereof). Suppose that the organism's choice rule is 'maximize expected utility, relative to my current belief state' and that its update rule is Bayesian conditionalization. Then, it follows that the organism will implement an evolutionarily optimal policy as characterized in section 4; this can be seen by direct inspection of expression (3).

What could license the assumptions that the organism's initial belief state is correct (i.e., matches the true distribution on the states) and that its utility function matches the true payoff function? A short answer is 'evolution by natural selection'. Organisms often have information about the environment 'preprogrammed' into their genome, as a result of natural selection, and are often equipped with 'appropriate' desires for surviving and reproducing (e.g., many animals are born with an innate ability to determine who their close relatives are and innate preferences for some food items over others). From an adaptationist perspective, therefore, the assumptions in question do not seem unreasonable.

https://doi.org/10.1086/674058 Published online by Cambridge University Press

Given these assumptions, our argument shows that an organism will implement an evolutionarily optimal policy if it chooses between actions in accordance with expected utility maximization relative to its current belief state and always updates its belief state by the rule of conditionalization. Therefore, two standard principles of Bayesian rationality fall out of the requirement of evolutionary optimality. This suggests that, other things being equal, natural selection should lead to organisms that satisfy these Bayesian principles.

This conclusion needs qualifying in a number of ways. First, the inference from 'is evolutionarily optimal' to 'is likely to have evolved' can obviously be questioned; this is a well-known issue in the literature on adaptationism. Second, defining optimality in terms of maximal expected reproductive output raises some tricky theoretical issues discussed in the next section. Third, our argument does not show that the only way for an organism to implement an evolutionarily optimal policy is to operate with internal belief-like and desire-like representations that combine and update in accordance with Bayesian principles. This last point merits some brief discussion.

It is well known from the work of Gerd Gigerenzer that adaptive behavior may be produced by simple heuristics, which are computationally less demanding than optimality calculations (e.g., Gigerenzer, Todd, and Pachur 2011). This general point applies to our problem of choosing an optimal policy. Even if there are many states of nature and signals, an organism might still be able to implement an optimal policy by a simple heuristic such as 'if signal x is received, run away; otherwise, stay', depending on the payoff function. And in fact, many simple animals exhibit behavioral plasticity in response to environmental cues but presumably do not operate with a belief-like representation of the environment at all. Of course in a purely 'as if' sense, these animals are behaving like Bayesian agents so long as they are implementing an optimal policy, as our argument shows. But we should not blur the distinction between organisms that implement an optimal policy by using a heuristic and ones that operate with internal belief-like and desire-like representations and are thus capable of satisfying Bayesian principles in more than an 'as if' sense. An important question is when evolution will lead to organisms of each type. This question is not addressed by the foregoing analysis; see Sterelny (2003) for a sketch of a possible answer.

6. Relation to the Brown/Maher Argument. It is worth relating our evolutionary argument to a related argument found in the rational choice literature. Both Brown (1976) and Maher (1992) try to justify Bayesian conditionalization by arguing that it follows from the injunction to maximize expected utility. They consider an agent with a prior subjective probability distribution over a set of states of nature, a set of possible actions, and a

https://doi.org/10.1086/674058 Published online by Cambridge University Press

utility function defined on the (state, action) pairs. The agent receives a signal, updates to a new probability distribution, then chooses an action that maximizes expected utility relative to this new distribution. Brown and Maher then show that in order to maximize expected payoff with respect to the agent's prior distribution, their update rule should be conditionalization. They interpret this to mean that rationality requires an agent to update in a Bayesian way.

Formally, the Brown/Maher result is a special case of our own. For the Brown/Maher argument identifies a particular way of implementing an optimal policy, for an agent who uses an update rule to go from belief state to belief state and a choice rule to go from belief state and utility function to choice of action. Brown and Maher assume that the choice rule is expected utility maximization relative to the updated state and, then, show that the update rule should be conditionalization. Our result is more general in that it does not assume from the outset that the organism implements its policy by using an (update rule, choice rule) combination at all. However, it follows directly from our result that if an organism implements its policy this way, and if its choice rule is 'maximize expected payoff relative to the updated state', then its update rule should be conditionalization—which is the Brown/Maher result. This follows directly from the characterization of an optimal policy in expression (3).

To see that our result is stronger than the Brown/Maher result, notice that theirs leaves open the possibility that an agent could earn a higher expected payoff, relative to the prior distribution p, by using an update rule other than conditionalization and a choice-of-action rule other than maximize expected payoff. This possibility is not ruled out by the Brown/Maher argument since they take for granted that the choice-of-action rule is maximize expected payoff. However, our result rules the possibility out. Since the combination (conditionalization, expected payoff maximization) implements an optimal policy, it follows that no other combination can do better, even if choice-of-action rules other than expected payoff maximization are permitted.

7. A Qualification. The foregoing argument assumes that expected reproductive output is the right criterion of evolutionary success. But in fact this is not always the case. It is well known that in certain situations, the expected performance of a strategy (or phenotype) is not the sole determinant of whether it will evolve; variability in performance can also matter (Gillespie 1977; Frank and Slatkin 1990; Seger and Brockmann 2002). How does this complication affect the foregoing argument?

Expected reproductive output is an appropriate criterion of evolutionary success (or 'definition of fitness') when two conditions are met. First, the population must be large, and second, the probability distribution over

https://doi.org/10.1086/674058 Published online by Cambridge University Press

states of nature that any organism faces must be independent across the members of the population. These conditions jointly ensure that the average reproductive success of the cohort of organisms using a given strategy will be very close to that strategy's expected success. The large population condition will often be satisfied, but the independence condition is more problematic.

In the predator example of section 3, the independence condition seems reasonable. The probability that organism A encounters a snake while foraging may well be independent of the probability that organism B does; risks of this sort are sometimes called 'idiosyncratic'. But in other cases risks will be aggregate. The weather is an obvious example, as it affects many organisms at once. If there is a 5% chance of a harsh winter, this risk will obviously not be independent across all members of a biological population. Real organisms probably face both idiosyncratic and aggregate risks, in varying combinations depending on the context.

Where risks are aggregate, the strategy with the highest expected reproductive output need not be evolutionarily optimal (McNamara 1995; Robson 1996). The variance in output matters too, and selection will penalize strategies with a high variance. In such a case, our simple argument for the evolutionary optimality of Bayesian updating does not apply. This suggests that in theory, organisms whose behavior is mediated by internal belief-like and desire-like states might do best to use Bayesian updating some of the time but not always. In a circumstance in which risks are idiosyncratic, as in our predator example, an organism should incorporate new information by Bayesian updating, assuming that its choice-of-action rule is 'maximize expected payoff in the updated state'. But where risks are aggregate, as in the weather example, a non-Bayesian update rule might be evolutionarily superior. However, to implement this selectively Bayesian strategy would require being able to distinguish aggregate from idiosyncratic risks. It seems unlikely that organisms can do this, but it is an empirical question.

REFERENCES

Brown, Peter M. 1976. "Discussion: Conditionalization and Expected Utility." *Philosophy of Science* 43:415–19.

Christensen, David. 1991. "Clever Bookies and Coherent Beliefs." *Philosophical Review* 2: 229–47.

Cooper, William S. 2003. *The Evolution of Reason*. Cambridge: Cambridge University Press.

Frank, Steven A., and Montgomery Slatkin. 1990. "Evolution in Variable Environments." *American Naturalist* 136 (2): 244–60.

Gigerenzer, Gerd, Peter Todd, and Thorsten Pachur, eds. 2011. *Heuristics: The Foundations of Adaptive Behaviour*. Oxford: Oxford University Press.

Gillespie, John. 1977. "Natural Selection for Variances in Offspring Number: A New Evolutionary Principle." *American Naturalist* 111:1010–14.

https://doi.org/10.1086/674058 Published online by Cambridge University Press

Greaves, Hilary, and David Wallace. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* 115:607–32.
Houston, Alasdair I., John M. McNamara, and Mark D. Steer. 2007. "Do We Expect Natural Selection to Produce Rational Behaviour?" *Philosophical Transactions of the Royal Society* B 362:1531–43.
Leitgeb, Hannes, and Richard Pettigrew. 2010. "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77:236–72.
Lewis, David. 1999. "Why Conditionalize?" In *Papers in Metaphysics and Epistemology*, 403–7. Cambridge: Cambridge University Press.
Maher, Patrick. 1992. "Diachronic Rationality." *Philosophy of Science* 59:120–41.
McNamara, John M. 1995. "Implicit Frequency-Dependence and Kin Selection in Fluctuating Environments." *Evolutionary Ecology* 9:185–203.
McNamara, John M., Richard F. Green, and Ola Olsson. 2006. "Bayes' Theorem and Its Applications in Animal Behaviour." *Oikos* 112:243–51.
Robson, Arthur. 1996. "A Biological Basis for Expected and Non-expected Utility." *Journal of Economic Theory* 68:397–424.
Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: Wiley.
Seger, Jon, and H. Jane Brockmann. 2002. "What Is Bet-Hedging?" *Oxford Surveys in Evolutionary Biology* 4:182–211.
Skyrms, Brian. 1995. *Evolution of the Social Contract*. Cambridge: Cambridge University Press.
Sterelny, Kim. 2003. *Thought in a Hostile World*. Oxford: Blackwell.
van Fraassen, Bas. 1989. *Laws and Symmetry*. Oxford: Oxford University Press.

https://doi.org/10.1086/674058 Published online by Cambridge University Press