Do we expect natural selection to produce rational behaviour?

Alasdair I. Houston[*]**, John M. McNamara and Mark D. Steer**

Centre for Behavioural Biology, University of Bristol, Bristol BS8 1RJ, UK

We expect that natural selection should result in behavioural rules which perform well; however, animals (including humans) sometimes make bad decisions. Researchers account for these with a variety of explanations; we concentrate on two of them. One explanation is that the outcome is a side effect; what matters is how a rule performs (in terms of reproductive success). Several rules may perform well in the environment in which they have evolved, but their performance may differ in a 'new' environment (e.g. the laboratory). Some rules may perform very badly in this environment. We use the debate about whether animals follow the matching law rather than maximizing their gains as an illustration. Another possibility is that we were wrong about what is optimal. Here, the general idea is that the setting in which optimal decisions are investigated is too simple and may not include elements that add extra degrees of freedom to the situation.

Keywords: transitivity; revenge effects; rate maximization; trade-off; predation; delay reduction

1. INTRODUCTION

In this paper, we are concerned with modelling the action of natural selection. Selecting the best action when making a decision can be of great importance to an individual's fitness; we often term making the right choice as being 'rational'. Ultimately, however, decision-making processes are products of an individual's evolutionary history, endowed via natural selection. In this paper, we are concerned with whether we should expect rational behaviour to be a product of naturally selected systems. In a general sense, rationality involves thinking and behaving reasonably and logically, but the term holds different meanings for researchers in different intellectual fields. The meaning and implications of the term 'rationality' have been discussed at great length. We do not intend to try and review all of these here, but use the categorizations of Kacelnik (2006) as a guide (for further introductions to the debate, see Wilson 1974; Moser 1990; Manktelow & Over 1993). Kacelnik (2006) adeptly introduces and summarizes three categorizations, representing the different disciplines for which rationality has been of central interest: philosophy and psychology (PP-rationality); economics (E-rationality); and behavioural ecology/evolutionary biology (B-rationality).

Psychologists have traditionally been more interested in the internal mechanisms of behavioural processes, rather than the behavioural outcomes *per se*. Rational behaviour is distinguished from irrational as a function of the process by which the behaviour became manifest, not by the behaviour itself (akin to Simon's (1976) procedural rationality). An important consequence is that PP-rationality is not understood in terms of observable behaviours, but in terms of internally consistent thoughts and beliefs (Kacelnik 2006). As a result, it is difficult to carry out experimental analyses of PP-rationality, since it involves examination of the cognitive processes involved in producing behaviour. Human subjects can be questioned about their reasons for choosing particular courses of action, enabling a limited level of investigation into beliefs; such investigation is virtually impossible in the study of non-human behaviour (Kacelnik 2006).

In contrast to PP-rationality, E-rationality is predominantly a goal-led concept, within certain bounds. The goal is the maximization of expected utility. E-rationality can therefore be used to predict patterns of observable behaviour (Kacelnik 2006). For the studies of human rationality, utility is most often taken as a financial gain, although there are some notable exceptions (e.g. Silberberg *et al.* 1991); in non-humans, utility is most often assumed to be linked to food acquisition, but could involve access to water, mates or conspecifics. However, the plasticity of utility as a term brings its own problems. A forager that seems to be failing to maximize the obvious utility may not be behaving irrationally but maximizing a different utility. As long as observed choices can be shown to maximize some form of utility, no matter how bizarre, then a decision maker can be classed as acting rationally. E-rationality suffers from another problem when its underlying axioms do not hold true. Violations of the axioms of transitivity and independence have been repeatedly reported from a variety of literatures (e.g. Busemeyer & Townsend 1993; Rieskamp *et al.* 2006; and references therein). Some examples are discussed in more detail later. A rigorous logician might argue that violations of the axioms of E-rationality can only mean one of two things: either that animals, including humans, cannot be rational or E-rationality cannot be a reasonable description of behaviour. In response, it might be argued that E-rationality, on the whole, can provide a good account of behaviour and the

* Author for correspondence (a.i.houston@bristol.ac.uk).

underlying assumptions are violated only in extreme circumstances. It is indeed true that economic concepts of rationality have proved to be effective and useful in predicting behaviour, but the recent boom in experimental economics (e.g. Hey 2002; Kahneman 2003; Smith 2003), which studies how economic agents actually behave without disregarding deviations from how they ought to behave, illustrates that it has not been totally successful (Hammerstein & Hagen 2005; Kacelnik 2006).

The differences between PP- and E-rationality are relatively straightforward. PP-rationality is a process-based concept dealing with predominantly internal beliefs and not their outcome. On the other hand, E-rationality considers the outcome to be of primary importance, in that utility should be maximized, but the process by which this occurs is not taken to be of great interest. B-rationality, springing mainly from the evolutionary literature, has yet another approach.

Steer & Cuthill (2003) proposed that a direct analogy can be drawn between optimal behaviour in animals and rational behaviour in humans, and that lessons learned by the study of the former can be applied to the latter. Studies in both biological and psychological literatures have tested the premises of E-rationality in non-humans (e.g. Navarick & Fantino 1972, 1974, 1975; Shafir 1994; Hurly & Oseen 1999; Waite 2001*a*, *b*; Bateson 2002; Bateson *et al.* 2002, 2003; Shafir *et al.* 2002; Waite & Passino 2006). In one sense, B-rationality, also known as ecological rationality (Todd & Gigerenzer 2000; Stephens *et al.* 2004; Hutchinson & Gigerenzer 2005), can be thought of as a subset of E-rationality, in that it defines a desired outcome, with utility being replaced by fitness, which is a more specific concept. An important difference between fitness and utility is that fitness functions can be measured, in terms of reproductive success, independently of the decisions an agent makes, whereas utility functions are derived from the decisions themselves and therefore are not independent of the choice procedure (Luce & Raiffa 1957; Houston & Staddon 1981; Kacelnik 2006). The processes by which an animal reaches a decision are not of primary importance to B-rationality, again similar to E-rationality. However, there is an added caveat: B-rationality assumes that agents are products of naturally selected processes which have shaped the cognitive and emotional machinery of the decision-maker to behave in a manner such as to maximize fitness. We cannot expect natural selection, having no foresight, to shape organisms to act rationally in all circumstances, but only in those circumstances which it encounters in its natural setting. Therefore, B-rational behaviour (i.e. fitness-maximizing behaviour) might not appear when animals are placed in a novel context. As an example, take an anecdote concerning the ultimate father of B-rationality, Charles Darwin. While on the Galapagos, Darwin noted that a frightened marine iguana (*Amblyrhynchus cristatus*) could not be induced to enter the ocean by any means other than picking the animal up and tossing it into the waves (something he did a number of times to one unfortunate individual). He wrote 'perhaps this singular piece of apparent stupidity may be accounted for by the circumstance, that this reptile has no enemy whatever on shore, whereas at sea it must often fall a prey to the numerous sharks. Hence, probably urged by a fixed and hereditary instinct that the shore is its place of safety, whatever the emergency may be, it there takes its refuge.' (Darwin 1839, ch. 17). Only when the context of the decision changed with the appearance of a land-based agitator did the iguana's behaviour—heading for land given any danger—appear to be irrational. It is clear to see then that B-rationality can be consistent with the use of simple rules of thumb (heuristics; Todd & Gigerenzer 2000) to determine behaviour (McNamara & Houston 1980).

To summarize, PP-rationality focuses on how decisions or beliefs are arrived at, but not necessarily on what the decisions or beliefs actually are. Conversely, the focal point of E-rationality is the decision itself, not the process by which it is achieved. E-rationality assumes that an agent will attempt to act in such a way as to maximize utility, utility being an undefined entity. B-rationality, similar to E-rationality, is also most concerned with the endpoint of a decision-making process, but assumes that an animal will maximize fitness when in a relevant context (Kacelnik 2006). From here on, we concentrate on the interplay of E- and B-rationality.

E-rational preferences are generally assumed to obey a series of conditions, including the following.

(i) *Transitivity.* In its simplest form this states that preferences are hierarchical, so if option *a* is preferred to option *b* and option *b* preferred to option *c*, then *a* will be preferred to *c*.

(ii) *Independence from irrelevant alternatives* (I.I.A.; Arrow 1951; Tversky & Simonson 1993). The basic idea is that relative preference for one option over another is unaffected by adding or removing options from the choice set (see also Luce 1959, 1977; Luce & Suppes 1965; Rieskamp *et al.* 2006).

If either of these conditions is violated, behaviour is classed as irrational.

As we will show, existing models of choice may predict seemingly irrational behaviours in particular circumstances. Broadly speaking, models of choice are descriptive (they describe the observed behaviour) or normative (they specify the behaviour that ought to be observed). Normative models are based on the evaluation of behaviour in terms of some measure (e.g. money or reproductive success). These approaches are linked; if rules have been shaped by natural selection, then rules that provide a good description should also make sense in terms of performance. In this paper, we review examples of behaviour that at first sight do not conform to what we might expect from a rational decision maker. We also present new results on a decision principle known as the delay-reduction hypothesis (DRH) and bring out general patterns in the behaviour of humans and other animals.

2. DESCRIPTIVE MODELS

The wealth of descriptive models that have been put forward is too great for us to review them all here.

Instead, we focus on a few descriptive models from operant psychology that have been influential in the study of choice behaviour.

(a) *Matching*

The matching law (Herrnstein 1961, 1970) emerged as a description of how animals in the laboratory choose between options that provide food. The matching law states that, for two options, the ratio of responses which a decision-maker makes on the options equals the ratio of rewards that it has previously received from the options, i.e.

$$\frac{B_1}{B_2} = \frac{R_1}{R_2}, \tag{2.1}$$

where R_i is the number of rewards previously received from option i and B_i is either the number of responses previously made or the amount of time previously spent responding on that option. R_i/B_i is the local rate from option i, so the matching law means that local rates are equal. A proportional form of the law is often used, which is as follows:

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2}. \tag{2.2}$$

Baum (1974) gave the following generalization of the matching law:

$$\frac{B_1}{B_2} = b \left(\frac{R_1}{R_2} \right)^s, \tag{2.3}$$

where s and b are the fitted parameters often referred to as sensitivity and bias, respectively. Houston and colleagues (Houston & McNamara 1981; Houston & Sumida 1987) caution that these parameters should be treated simply as fitted parameters, since their relevance to actual choice mechanisms is not clear. If both s and b equals 1, then basic matching holds, since Herrnstein's original formulation of the matching law adequately describes the data. The generalized matching law, however, effectively allows many behavioural patterns that deviate from basic matching to be explained in terms of generalized matching, but only by fitting the values of s and b *a posteriori*.

Since its inception, the matching law, especially in its generalized form, has been a successful and popular tool for describing behaviour (e.g. Myerson & Hale 1984; McDowell 1989; Pierce & Epling 1995; Spiga *et al.* 2005). It has been used to explain behaviours as diverse as wagtail foraging (Houston 1986*a*) and self-harm in humans (McDowell 1981; however, see Fuqua 1984 who highlights a range of inconsistencies between controlled investigations of matching behaviour and tests of matching in more applied settings). The relationship between matching and neurophysiology has also been discussed (e.g. Sugrue *et al.* 2004; Soltani & Wang 2006). However, it is important to note that basic matching (equation (2.1) or (2.2)) may not uniquely specify behaviour; in some settings, matching can be produced by a wide range of different behavioural allocations (Houston & McNamara 1981).

Many studies of matching behaviour have investigated the behaviour of animals (including humans) on concurrent variable interval (VI) schedules. These schedules are often used in experimental psychology to test choice behaviour; well over 50 studies have been published in the last 10 years, which have used these schedules. They supply a reward (or stimulus of some sort) following a subject's first response after a given, but variable, delay has elapsed. For example, a pigeon might be confronted with an illuminated disc (or 'key') that it can peck. During an initial random delay (the variable interval), any pecks that the pigeon makes on the key are unrewarded. However, once the delay has elapsed, the first peck that the pigeon makes on the key results in a reward being delivered (see Ferster & Skinner 1957 for further information). If two VI schedules are available to a subject at any one time, then this is termed a concurrent VI–VI procedure.

Optimal foraging theory is a normative approach that attempts to explain behaviour in terms of the maximization of fitness (see Stephens & Krebs 1986 for a review). One simple assumption is that maximizing the rate of energetic gain will maximize fitness. Several papers have discussed the relationship between the matching law and rate maximization. It has been shown that maximizing the rate of gain does not necessarily result in matching (e.g. Heyman & Luce 1979; Houston & McNamara 1981; Houston 1983; Heyman & Herrnstein 1986). Given that matching does not necessarily specify behaviour uniquely, it is not possible to say whether matching behaviour maximizes rate of gain. On concurrent VI schedules, an infinite number of behavioural allocations can satisfy matching. Some of these allocations will give rates that are close to optimal (Houston & McNamara 1981). When faced with a VI schedule and a schedule that has a constant probability of giving a reward, matching results in a rate of reward that is well below the optimal (Houston 1983; Heyman & Herrnstein 1986). Behaviour on such schedules can be described by the generalized matching law (Heyman & Herrnstein 1986). This behaviour and the resulting loss in reward rate can be viewed as side effects of a decision rule that evolved in other circumstances.

VI schedules often use a negative exponential distribution of times between rewards. This means that whether a response is rewarded or not gives no information about the time until the next reward becomes available. The form of the best strategy when choosing between two such schedules is to repeat a cycle comprising a fixed time t_1 on side 1 and a fixed time t_2 on side 2 (Houston & McNamara 1981). In other words, if the animal knows the parameters and is sure that they will not change, it should ignore rewards and get an accurate clock so that it can measure the optimal times t_1 and t_2. Houston & McNamara (1981) found an exact solution to the problem of maximizing rate of gain, given a choice between two VI schedules when the mean interval of each schedule is known (see also Belinsky *et al.* 2004). This is not realistic. If we wish to understand the evolution of foraging behaviour, we should be looking for rules that perform well under the range of conditions that an animal is likely to experience (cf. Seth 2007). It should not be assumed that the animal has full knowledge of current conditions. Instead, the animal both learns about and exploits its environment. There are three general features of foraging environments that are relevant here.

(i) *Rewards may give information about future rewards.* This possibility has been thoroughly investigated in the context of how long to stay in a patch that contains a random number of food items (e.g. Iwasa *et al.* 1981; McNamara 1982; McNamara & Houston 1987*c*).

(ii) *The environment may contain other foragers.* A rule that works well for an isolated forager might not work well when that animal has to compete with others. Similarly, a forager that uses a rule that is successful in group situations might perform poorly when in isolation (Seth 2001, 2007).

(iii) *The environment may change.* If a set of environmental parameters is constant, then it is often possible to evolve fixed optimal behaviours. However, these behaviours will become suboptimal, given any change in the environment (McNamara & Houston 1985, 1987*b*; McNamara 1996; Dall *et al.* 1999).

To cope with these aspects of the real world, what is needed is an approach that is based on rules that use information from rewards obtained to decide between options. A process called melioration uses previous information and results in outcomes that satisfy the matching law. The idea behind melioration is that an animal increases its allocation to the alternative that gives it the highest local rate. At the stable outcome with both options chosen, local rates are equal, i.e. matching holds. This is really a framework rather than a detailed model; there are lots of ways in which melioration can be implemented. There are also rules that result in matching without using the principle of melioration (e.g. Harley 1981; Houston & Sumida 1987; Seth 1999).

(b) *The delay-reduction hypothesis*
Like the matching law, the DRH (Fantino *et al.* 1993) was developed as a description of choice behaviour in the laboratory. The DRH can be used to predict choice on what is known as the concurrent chains procedure. In the simplest case, an animal can respond to one of two alternatives, known as initial links. Each alternative has an associated VI schedule but, in contrast to a standard concurrent VI–VI procedure in which the VIs provide rewards, here, the VIs provide access to terminal links that result in a reward after a delay has elapsed. This access is indicated by a cue that signals the availability of a reward after a certain delay. During the initial link phase, the animal can choose between the two alternatives, but once a terminal link becomes available, the animal must wait until this link ends in a reward. After this has occurred, the animal can again choose between the initial links. There is some resemblance between the chains procedure and an animal that is searching for two food types at once. From time to time, it encounters food items that result in energy after the handling time has elapsed. These items can be thought of as being analogous to the terminal links. Let T be the overall average time to a reward on the concurrent chains procedure. This time depends on both the initial links and the terminal links. To illustrate, assume that each initial link is a VI, with the time to its terminal link having an exponential distribution with a mean of 60 s.

Then, the average time for the first terminal link to become available is 30 s. If one terminal link has a delay of 10 s and the other has a delay of 30 s, then the average time on terminal links is 20 s and hence $T=50$ s (see equation (2.5) for the general equation for T). The start of the terminal link on side i means that food will be available after a delay D_i. Thus, the start of the terminal link is associated with a reduction in the expected delay to reward of $T-D_i$. Define ρ to be the proportion of responses to alternative 1, i.e. the number of responses made on the initial link for side 1 divided by the total number of responses made on both the initial links. The DRH for the two alternatives states that

$$\rho = \frac{T - D_1}{T - D_1 + T - D_2}. \tag{2.4}$$

This is the equation suggested by Fantino (1969). (For modifications in the case of unequal initial links, see Squires & Fantino 1971; Fantino & Davison 1983.)

Fantino & Dunn (1983) point out that the DRH predicts the violation of the principle of independence from irrelevant alternatives (I.I.A.). Fantino & Dunn (1983) and Mazur (2000) found that adding a third option in a concurrent chains procedure could change the preference of pigeons for the initial pair of options. We now give a formal analysis based on the version of the DRH given in equation (2.4). Consider a general case with n initial links, each with an exponential VI schedule with mean interval I. Then, the mean wait until a terminal link first becomes available is $W_n=I/n$. The DRH states that an animal's preference for an alternative is given by the relative reduction in delay to reward associated with the alternative. If initial link i leads to a terminal link with delay to reward D_i, then the overall time to reward is

$$T_n = W_n + \frac{1}{n} \sum_i D_i. \tag{2.5}$$

Now, compare the ratio of allocations to options 1 and 2 with and without option 3 being present. In the case of just two options being available,

$$\frac{R_1}{R_2} = \frac{T_2 - D_1}{T_2 - D_2} = \frac{I + D_2 - D_1}{I + D_1 - D_2}.$$

When a third option is added, the allocation is

$$\frac{R_1}{R_2} = \frac{T_3 - D_1}{T_3 - D_2} = \frac{I + D_2 + D_3 - 2D_1}{I + D_1 + D_3 - 2D_2}.$$

It follows from these equations that adding a third alternative can change the allocation of responses to option 1 relative to option 2. In other words, we have a violation of I.I.A. The relative allocation may either increase or decrease; the critical value for a third alternative to produce no change is

$$\hat{D}_3 = \frac{I + D_1 + D_2}{2} = T_2.$$

Otherwise, adding a third alternative does have an effect.

DRH and optimal foraging theory
We now explore the relationship between the DRH (a descriptive account) and the optimal foraging theory (a normative account) and present a new result relating

the DRH to the costs of deviating from optimality. Consider a foraging animal that searches for food and encounters two types of prey item. The DRH equation for two options (equation (2.4)) applies to cases in which the animal responds on both the initial links. This corresponds to the region of parameter space in which the maximization of rate of energetic gain predicts that both the prey types should be accepted (Fantino & Abarca 1985; Houston 1991). This means that the optimal rate of energetic gain, γ, is the rate resulting from taking both the types. Assuming equal energy content (which can be set equal to 1 without loss of generality) and denoting the overall time to reward by T,

$$\gamma = \frac{1}{T}.$$

Now, consider an animal that encounters items that can differ in energy and handling time. An item of type i has energy e_i and handling time D_i. McNamara & Houston (1987a) show that the energetic value of accepting a type i item is

$$H_i = e_i - \gamma D_i.$$

As Houston & McNamara (1999) point out, e_i is the energy gained by accepting a type i item and γD_i is the energy that could have been obtained by foraging at rate γ for time D_i rather than spending this time handling the item. Thus, H_i is the energetic value of accepting the item. It is optimal to accept all item types for which $H > 0$, i.e. for which $e/D > \gamma$ (Houston & McNamara 1999). As we have already said, the DRH is concerned with parameter values for which both 'types' should be accepted. This means that H_i is positive for both the alternatives and is the energy lost if the type is rejected rather than accepted. In other words, H_i is the cost associated with making the error of rejecting type i.

When $e_1 = e_2 = 1$,

$$H_i = 1 - \frac{D_i}{T}. \tag{2.6}$$

Animals typically make errors of decision and errors are more probable if the cost is low (Houston 1987, 1997; McNamara & Houston 1987a). One possible simple equation that captures this property is

$$\rho = \frac{H_1}{H_1 + H_2},$$

where ρ is the probability of choosing option 1. If we use equation (2.6) to substitute for H_1 and H_2 in this equation, then we obtain equation (2.4).

Thus, we have shown that the DRH is linked to optimality and the costs of making errors. This might suggest that it is a good rule—even though it was proposed as, and is still primarily, a descriptive model of choice, it has a normative basis.

So far, we have discussed the DRH when all rewards have the same magnitude. Data from experiments in which terminal links differ in delay and reward magnitude violate a form of transitivity known as strong stochastic transitivity (Navarick & Fantino 1972, 1974, 1975). Houston (1991) shows that a generalization of the DRH to include different reward magnitudes can produce such violations. Houston *et al.* (2005) extend the analysis to include terminal links in which the delay

to reward is variable and show that stochastic transitivity can still be violated when reward magnitudes are constant, but the terminal link durations are variable.

3. NORMATIVE MODELS
We expect rules to be 'good', i.e. natural selection has resulted in rules that perform well. But animals (including humans) sometimes make bad decisions ('humans are not rational'—see Sutherland 1994 for a review). Such findings have been explained in a number of ways; we concentrate on two of these explanations.

(a) *The outcome is a side effect*
What matters is how a rule performs (in terms of reproductive success). Several rules may perform well in the environment in which they have evolved. Their performance may differ in a 'new' environment (e.g. the laboratory); some rules may perform very badly in this environment. The debate about whether animals follow the matching law rather than maximizing their rate of energetic gain can be used as an illustration. It is possible that what has been favoured by natural selection is a rule that performs well in environments that do not resemble experimental procedures, and hence, when a forager is presented with such procedures, the behavioural outcome is not what we might expect from an E-rational standpoint. It is important to distinguish between rules and outcomes when considering whether a particular behaviour is rational or not. We may well be able to liken some behavioural outcomes to spandrels (*sensu* Gould & Lewontin 1979), in that they are non-selected by-products of a decision-making mechanism (rule). Take risk sensitivity for example; although there are various normative explanations for the appearance of risk-sensitive responses (e.g. McNamara & Houston 1992; Houston & McNamara 1999), it has also been argued that risk sensitivity can appear as a side effect of a forager using simple learning rules (e.g. Kacelnik & Bateson 1996; March 1996). Arkes & Ayton (1999) suggest that some errors in human reasoning result from overgeneralizing a rule that is reasonable in many contexts. (For a general discussion of rules and side effects, see McNamara & Houston 1980.)

Decision framing is another important factor. There are lots of ways to provide an animal with a particular set of options, all of which have the same mathematical characterization, but which may result in systematic differences in behaviour. The general point is that the details of the experimental procedure can be important (e.g. Shettleworth 1989; Savastano & Fantino 1994; Heyman & Tanz 1995). Shettleworth & Jordan (1986) found that rats preferred receiving sunflower seeds in the husk and removing the husks themselves to simply waiting for a 'handling time' to elapse before being presented with a dehusked seed. Similarly, different time periods within an experimental situation seem to be treated with different degrees of importance by foragers; for example, inter-trial intervals are often found to be unimportant for guiding choice behaviour, whereas the delay between making a response and receiving food is extremely important (e.g. Kacelnik & Bateson 1996; Stephens *et al.* 2004). When seemingly irrational behaviours appear, especially in experimental situations,

we always need to ask whether it is the outcome of a well-adapted rule misfiring in a novel environment.

(b) *We were wrong about what is optimal*

The idea here is that the context in which optimal decisions are viewed is too simple and may ignore elements that add extra degrees of freedom to the situation. We now present a range of examples.

(i) *Uncontrolled variation in state*

Studies of human irrationality tend to concern one-shot decisions; there can therefore be no differential accumulation of rewards between treatments. This is not always the case for non-humans. Schuck-Paim *et al.* (2004) highlight several cases where analogies have been drawn between similar 'irrational' behaviours in humans and non-humans; however, they claim that the underlying choice mechanisms are fundamentally different. Their findings hold considerable implications for the comparison of choice behaviour across species. Schuck-Paim *et al.* (2004) show that, at least in some cases, seemingly irrational behaviour in animals can be explained purely as a function of state-dependent preferences. As examples, we discuss work on grey jays (Waite 2001*a*) and rufous hummingbirds (Bateson *et al.* 2002).

Grey jays (*Perisoreus canadensis*) were trained in one of two contexts, both involving choosing between two foraging patches. The patches consisted of a tube in which raisins were placed at different distances from the entrance; increased distance was assumed to correlate with increased perceived predation risk. In context A, the jays were offered a series of choices between one and three raisins placed 0.5 m along separate tubes. Birds in context B were offered a series of choices between two identical options: two tubes with a single raisin placed 0.5 m from the entrance. All birds were subsequently offered a choice between three raisins placed 0.7 m along a tube and one raisin just 0.3 m from the entrance of a second tube. Preference for the larger, but riskier, reward was higher among birds which had been trained in context B. The findings were taken as indicating departures from value maximization as a result of cognitive biases arising in consequence of the choice context. The results were seen as mirroring framing effects (Tversky & Kahneman 1981), in that the same decision was presented to different individuals but in a different scenario, the scenario affecting the final decision. Waite (2001*a*) compared the results with that of the trade-off contrast hypothesis (Tversky & Simonson 1993). This predicts that an individual will be more likely to choose a low-quality, cheap item over a high-quality, expensive one if the individual has already experienced a choice between items of similar quality, but with a smaller difference in cost. However, as Schuck-Paim *et al.* (2004) pointed out, it was not just the previous context in which decisions had been made that differed between the two treatment groups. During the initial phase of the experiment, the jays in context A had gained more than twice as many rewards as those individuals in context B. The problem therefore can be thought of in terms of energy–predation trade-offs the kind of which have been extensively discussed in behavioural ecological systems (e.g. Houston *et al.* 1993; Houston & McNamara 1999; Cuthill *et al.* 2000).

Many species of animal can only increase their rate of energetic gain by also increasing their probability of being killed by a predator (see Lima 1998 for a review). To predict choice, it is necessary to know the value of energy (gain from foraging) and the value of life (lost if killed) (Houston & McNamara 1988, 1989). These will typically depend on the animal's state. In general, an animal should accept a risk in order to obtain energy when reserves are low (Houston & McNamara 1988; McNamara 1990; Clark 1994). The lower energetic state of the individuals in context B means that they should have been more prepared to take the greater risk to achieve the higher pay-off than the context-A individuals that were more sated. Similarly, seemingly irrational behaviour was reported from rufous hummingbirds (*Selasphorus rufus*; Bateson *et al.* 2002). The hummingbirds changed their relative preferences for two options in the absence or presence of a decoy option, which provided a lower rate of gain than either of the other two options. Once again, however, the rate of energy gain differed between the two conditions, which could well have led to alterations in choice behaviour consistent with rational theory (Schuck-Paim *et al.* 2004). Giving weight to their argument, Schuck-Paim and her colleagues showed experimentally how seemingly irrational decisions in European starlings (*Sturnus vulgaris*) disappeared when energetic rates were equalized across treatments.

(ii) *Future expectations*

The best action to choose at present depends on future expectations (McNamara & Houston 1986; Houston & McNamara 1999 and references therein). For example, whether or not an animal should take risks in terms of predation in order to obtain food depends on whether food is likely to be plentiful and easy to obtain in the future. Adding an option to the set of available options changes what is possible in the future, and hence can change future expectations, even if it is not optimal to choose the additional option now. Thus, even if the new option is not chosen when added, its presence can change the current optimal choice. We give two examples of this effect.

Errors. Suppose that there are errors in decision making, with costly errors being rare. Then, future gains depend not only on the preferred option, but also on other options that are mistakenly chosen. Adding a suboptimal option now will thus affect future expectations, because this option is likely to be wrongly chosen in the future. This means that the value of an option depends on the context (i.e. on the other options that are available). As a result, violations of transitivity may occur (Houston 1997) and violations of I.I.A. may also occur.

Possible future states. Schuck-Paim *et al.* 2004 show that uncontrolled variation in state can produce behaviour that appears to be irrational. Houston *et al.* (2007) show how state-dependent effects can produce apparently irrational behaviour even when an animal's choice is measured in the same state. Consider an animal choosing at discrete times between foraging options that differ in terms of expected energetic gain and risk of predation. The animal's state is its level of energy

reserves. The animal dies of starvation if its reserves fall to zero. Option A provides little food but has no associated risk of predation. Option B provides slightly better (but still not good) food and involves a risk of predation. Option C provides good food and involves the same risk of predation as option B. Thus, in this model, the animal is faced with options that differ in terms of energetic gain and risk of predation. Houston *et al.* (2007) consider three environments. In the first environment, options A and B are available. In the second environment, options B and C are available. In the third environment, options A and C are available. In each case, Houston *et al.* (2007) use dynamic programming to find the strategy that maximizes long-term survival (cf. McNamara & Houston 1990*b*). As we would expect, this strategy is state-dependent, i.e. the optimal decision depends on the level of energy reserves. When options A and B are available, it is optimal to choose A when reserves are very high and B otherwise. When options A and C are available, it is optimal to choose A unless reserves are low, in which case C is chosen. There is an intuitive explanation for these results. In this example, option A is safe but has a low yield. Thus, the option is used when energy reserves are high. When energy reserves are low, the animal should take risks in order to obtain a higher rate of energy gain. When option A is present with option C, because C has a higher yield than B the animal can afford to delay using this risky option until its reserves fall to a lower threshold than if B had been present rather than C. Finally, when options B and C are available, option C is always chosen because it yields more food than B and has the same predation risk. From these three results, we see that at intermediate levels of reserves, B is preferred to A when these are the two options available, C is preferred to B and A is preferred to C. In other words, transitivity is violated.

We have given two examples of violations of transitivity, one based on errors (Houston 1997) and the other based on state (Houston *et al.* 2007). The common principle is that all available options influence choice because they have an effect on future expectations. It is important to emphasize that these examples are based on the assumption that all options currently available will also be available in the future. Options are linked to the future either because the animal makes mistakes and hence may choose a suboptimal option or because stochastic changes in state may take it to a state in which an option that is not currently chosen should be used. If this sort of analysis is to be relevant, animals must expect some degree of persistence of options into the future. Whether we might predict animals to have this view will depend on the sort of environment in which they have evolved.

(iii) *There may be more freedom in behaviour than originally anticipated*
Activities often have positive and negative effects. As a result, an animal may be faced with a situation in which it has to trade-off these effects. We have mentioned above that getting food may expose an animal to predators. This may mean that there is a trade-off between energy and predation. We are now interested in the consequences of changing some aspect of an animal's environment. We draw attention to the contrast between results if behaviour is fixed and results if the animal is free to change its behaviour. The same basic principles can be seen in a range of examples from humans and other animals. Improvements in safety, such as the introduction of seatbelts or airbags in cars, provide an illustration. If the behaviour of road users does not change, then there should be a reduction in injury level. The changes may, however, lead to people changing their behaviour, typically by behaving in a more dangerous way. This change may be strong enough to result in an increased level of injury (for data and discussion, see Peltzman 1975; Keeler 1994; Peterson & Hoffer 1994). Similarly, the obvious response to traffic congestion—building more roads—may cause an increase in traffic (e.g. Noland & Lem 2002; Cervero 2003; Goodwin & Noland 2003) to the extent that traffic jams are worse than they were before. Thus, when animals are free to change their behaviour, the consequences may be an effect on performance that is the opposite of what was expected (and desired). A physiological example concerns whether a bit of dirt is a good thing. Dirt may have a negative direct effect on disease, but a positive indirect effect through changes in the immune system. The 'hygiene hypothesis' (Strachan 1989) states that improved hygiene (and reduced family size) have reduced the extent to which children are exposed to infectious agents. The result is a change in the immune system that renders it more likely to give rise to allergic responses such as asthma (see Yazdanbakhsh *et al.* 2002; Romagnani 2004; Christen & von Herrath 2005 for further discussions). Tenner (1996) calls effects like these *revenge effects*. In analysing such examples, two approaches can be adopted. In one, behaviour in response to the change is 'given' (i.e. we adopt a descriptive approach). An alternative is to derive the behavioural response from considerations of optimality (i.e. we adopt a normative approach). This second approach is adopted in models that derive the behaviour of humans from the maximization of utility. Blomquist (1986) and Janssen & Tenkink (1988) use utility maximization to derive the dependence of a driver's behaviour on a parameter that corresponds to the level of a safety measure. They show that the effect of an improvement in safety may be substantially reduced by changes in behaviour. We now demonstrate the normative approach in other contexts, starting with examples based on the trade-off between obtaining energy and avoiding predators.

There are two ways of looking at optimal response to a change: (i) change in optimal behaviour (e.g. driving speed in the model of Janssen & Tenkink (1988)) and (ii) change in some aspect of performance, such as mortality. We show that the relationship between various environmental factors and both the optimal response and the resulting performance may not always be obvious.

Changes in optimal behaviour. Consider the following example (see also McNamara & Houston 1994): a forager has to reach a critical size in order to reproduce. How should it respond to a permanent change in the predation level? The answer depends on how behaviour and predation interact to determine mortality. Assume that the animal has the choice of how hard it works to

obtain food. Denote this rate of work by u. The animal's rate of intake of food is proportional to u. Its predation rate M is given by

$$M(u) = m_0 N(u) + \mu,$$

where the function $N(u)$ is increasing and accelerating and determines how predation rate changes with u for a given density of predators (m_0) and μ is a background mortality rate. There is thus a trade-off between gaining food and predation risk. We denote the optimal rate of working for food by u^*.

If a change in predation is a result of an increase in m_0, u^* decreases. This is because there will be a marked increase in predation if foraging intensity is high; therefore, the best response is to adopt a less dangerous foraging behaviour. In contrast, if μ increases, u^* increases. This is because the same increase in predation rate is imposed on all foraging options. Therefore, the best response is to reduce the time exposed to predation by growing faster (i.e. by working harder for food). This shows that the effect of an increase in danger may result in either an increase or a decrease in how hard the animal works for food.

Changes in the performance. When animals can trade-off energetic gain against predation risk, the effect of a change in the environment may be counter-intuitive (e.g. McNamara & Houston 1990*b*; Abrams 1993). For example, an increase in food availability can lead to a decrease in food intake or an increase in starvation (McNamara & Houston 1987*d*; 1990*a*; see McNamara & Houston 1994 figs. 2 and 3; also Houston & McNamara 1999). The effects arise because the animals change their behaviour in adaptive ways (see McNamara & Houston 1994 for a review). We now give some examples of changes in the performance as a consequence of adaptive behaviour.

McNamara & Buchanan (2005) model a situation in which an animal is exposed to a stressor, such as cold or high predation risk, for a period of time. During this period, the animal can choose the level of available resources to direct against the stressor. The more the animal diverts to combating the stressor, the less probable the stressor is to kill the animal. However, diverting resources from essential maintenance reduces the condition of the animal. As condition decreases, the probability of death from disease increases. Thus, the animal faces a trade-off between dying from the stressor or from the effects of poor condition. McNamara & Buchanan (2005) find the allocation of resources to combating the stressor that maximizes the probability that the animal will survive. They show that if the likelihood of death from disease at a given level of condition is decreased, the animal allows condition to deteriorate much more. The result is to increase the likelihood that the animal will die from disease. In this model, much of the mortality from disease occurs during recovery of condition after the stressor disappears.

Failure to take account of the fact that behaviour is flexible may make it difficult to detect important costs. If we vary a factor, an animal may respond by changing its behaviour or morphology in a way that we have not anticipated. This makes it hard to detect

the direct effect of the change. We give an example based on the diving behaviour of the animals that hunt for food underwater and return to the surface to breathe (e.g. puffins, otters). A dive cycle starts with the animal at surface. The animal travels to the foraging area at a particular depth, forages there for a time t and then returns to the surface where it spends a time s gaining oxygen. The amount of oxygen gained is a decelerating function of time at the surface. The total time travelling between the surface and the foraging area is τ. A simple approach assumes that the animal should maximize the proportion of time spent in the foraging area, subject to the constraint that the diver balances its oxygen over the cycle (Kramer 1988; Houston & Carbone 1992). The rates of oxygen use are m_t and m_τ during foraging and travelling, respectively. If τ is increased with t fixed, then the oxygen constraint means that s is an accelerating function of time underwater. But if the animal is free to adopt the behaviour that maximizes the proportion of time spent in foraging, then as τ increases, s may be approximately proportional to time underwater (Houston & Carbone 1992; McNamara et al. 2001). Thus, the cost (the effect of a unit increase in τ on s as τ increases) is not apparent when the animal is able to adjust its time budget. Houston *et al.* (2003) investigate the behaviour of a diving animal that can catch only a single item when hunting underwater. They show that if the diver maximizes its rate of energetic gain while hunting for items of two types, then the success of a dive (i.e. the probability of returning to the surface with an item) is not a good indicator of the quality of the environment. For example, as the probability of finding the better type of item increases, the success of the dive may first increase, then decrease and then increase again.

It is often assumed that predation risk in birds depends on the fat load because heavy birds will be less agile. It is, however, hard to detect the effect of fat on predation because a bird may change both its behaviour and body composition. As fat loads increase, a bird may adopt safer foraging options. As a result, predation may decrease with mass (Welton & Houston 2001). Another response to an increased fat load is an increase in muscle mass, allowing greater agility and thus preventing an increase in predation.

These examples show that care is needed in choosing the variables that will be measured. If an important variable is not measured, then results may be misleading. For example, in the case of the diver, a better understanding of costs can be obtained if time at the surface is related to time travelling, τ, and time foraging, t, rather than to total time underwater, $\tau + t$.

(iv) Fluctuating environments and biased probabilities

It might seem obvious that natural selection should always result in organisms having an accurate view of the world. Models based on evolution in a certain kind of stochastic environment show that this is not the case. Chance acts on many scales in the natural world. At the finest scale, demographic stochasticity describes the good and bad luck that affects individual population members, independently of other population members. At the other extreme, environmental stochasticity

concerns fluctuations in the environment as a whole. These fluctuations, which might be due to weather or changes in population size, affect all population members in a similar manner.

Consider first the situation where there is demographic stochasticity, but no environmental stochasticity. Demographic stochasticity will affect the lifetime reproductive success (LRS) of individuals in the population. Let $p(x)$ be the probability that the LRS of a particular individual is x. Then, the mean LRS of this individual is

$$r = \sum_x xp(x),$$

where this mean is an average over demographic stochasticity. The quantity r is the standard fitness measure in this situation and tends to be maximized by the action of natural selection.

Now suppose that there is also environmental stochasticity. The LRS of an individual will then depend on both demographic good and bad luck and on the state of the environment. Let $p(x|s)$ denote the probability that the LRS of the individual is x when the environmental state is s. The mean LRS when the environmental state is s is thus

$$r(s) = \sum_x xp(x|s).$$

In this situation, the standard measure of fitness is the geometric mean, G, of $r(s)$, where the mean is an average over the environmental state s. Equivalently, fitness can be taken to be $g = \log G$, the logarithm of G. This fitness measure can be expressed as

$$g = \sum_s \log(r(s))f(s),$$

where $f(s)$ is the probability that the environmental state is s. The quantity g tends to be maximized by the action of natural selection.

Now suppose that the above population is at evolutionary stability, so that population members are maximizing the fitness measure g. Let $r^*(s)$ denote the expected LRS of population members when the environmental state is s. Then, it can be shown that population members are also maximizing

$$\sum_s r(s)f^*(s),$$

where f^* is a certain probability distribution of environmental states. Under this distribution, the probability, $f^*(s)$, of state s is proportional to $f(s)/r^*(s)$. As this formula shows, the distribution f^* distorts the true probability distribution f, giving extra weight to environmental states for which population members do badly and reducing the probability of environmental states for which population members do well. Thus, population members are maximizing their average LRS (averaged over environmental stochasticity), but the average is based on biased probabilities (McNamara 1995; cf. Haccou & Iwasa 1995; Sasaki & Ellner 1995). For the link between this approach and a general account of optimization under the action of natural selection, see Grafen (1999).

4. DISCUSSION

Cooper (2001) argues that evolution results in rational choice as summarized by the laws of logical thought. A limitation of Cooper's general line of argument is that it is based on choices made by rules that are optimal for particular conditions. We have stressed, however, that problems arise when animals are faced with novel environments (Darwin's iguana, see also McNamara & Houston 1980; Shettleworth 1985). Given that animals follow robust rules, side effects (often referred to as spandrels in evolutionary psychology, e.g. Buss *et al.* 1998; Hampton 2004) will be ubiquitous. A related point is that we cannot just determine optimal behaviour for the environment of the laboratory (Houston & McNamara 1989, 1999). An animal in the laboratory may be safe from starvation and predation, but it does not 'know' this. It presumably follows rules that evolved to cope with these threats and deal with competition and changes in the environment. Matching is not optimal in some of the procedures that are used in laboratory experiments; it may be that matching is a side effect of decision rules that perform well in a broader context (see also Seth 2007). We have shown that a version of the DRH can be related to the maximization of rate of energetic gain, given that errors occur but costly errors are rare. Our result suggests that although the DRH may not be strictly optimal, it is likely to be a good principle, given that errors occur. The interaction between options that is captured by the DRH can be understood in terms of decisions that are subject to error. From this view, some aspects of the DRH may have appeared to be irrational because we had a limited conception of optimality.

Previous work (e.g. Tversky & Simonson 1993) has presented models based on plausible psychological principles that can describe irrational behaviour. In this paper, we have attempted to construct links between descriptive and normative accounts. In addition to showing that the DRH emerges from optimal decision-making subject to errors, we have pointed out that intransitive choice can result from optimal behavioural mechanisms when decisions depend on state and options persist into the future. This result does not rely on uncontrolled variation in state. It emerges owing to the effect that options have on future expectations. This general principle deserves further investigation.

We have drawn attention to common themes that arise in the study of humans and other animals, but many analogies have not been explored. For example, the model of driving speed investigated by Janssen & Tenkink (1988) is analogous to models of optimal flight speed (e.g. Norberg 1981; Houston 1986b). Whether this resemblance is productive remains to be seen. An area in which a unified account might be useful is optimal defence. In the context of military history, we might be interested in how the builder of a castle should allocate resources to structures that improve the strength of the castle and to features that improve its appearance and hence the prestige of the builder, and consequently the number of descendants that he leaves. Analogous issues arise in several biological contexts, including the interactions between predators and prey (e.g. Abrams 1986; McNamara *et al.* 2005), the way in which plants defend themselves against herbivores (e.g. Adler &

Karban 1994; VanDam *et al.* 1996), the evolution of diseases and the defences against them (e.g. Frank 1996; Shudo & Iwasa 2001, 2004; Medley 2002; Day & Proulx 2004; van Boven & Weissing 2004) and the defence of a social insect colony against attack (e.g. Oster & Wilson 1978; Aoki & Kurosu 2004). Adler & Karban (1994) make the military analogy explicit and Jokela *et al.* (2000) present 'steps towards a unified defence theory', but we suspect that further work on a synthesis of these areas would be instructive.

We also think that an approach to decision making adopted by Fawcett & Johnstone (2003) could be extended. They investigate a model of optimal choice when an animal chooses between objects on the basis of more than one cue. These cues can differ in reliability of the information that they provide and in the cost of assessing them. This sort of approach may have broad implications for the understanding of apparently irrational behaviour.

The area that we have addressed is vast and our coverage has been highly selective; many issues have not been considered. For an entry to some of the topics that we have not discussed, see Bernardo & Welch (2001) and Robson (2002, 2003).

REFERENCES

Abrams, P. A. 1986 Adaptive responses of predators to prey and prey to predators: the failure of the arms race analogy. *Evolution* **40**, 1229–1247. (doi:10.2307/2408950)

Abrams, P. A. 1993 Why predation rate should not be proportional to predator density. *Ecology* **74**, 726–733. (doi:10.2307/1940800)

Adler, F. R. & Karban, R. 1994 Defended fortresses or moving targets—another model of inducible defenses inspired by military metaphors. *Am. Nat.* **144**, 813–832. (doi:10.1086/285708)

Aoki, S. & Kurosu, U. 2004 How many soldiers are optimal for an aphid colony? *J. Theor. Biol.* **230**, 313–317. (doi:10.1016/j.jtbi.2004.05.019)

Arkes, H. R. & Ayton, P. 1999 The sunk cost and Concorde effects: are humans less rational than lower animals? *Psychol. Bull.* **125**, 591–600. (doi:10.1037/0033-2909.125.5.591)

Arrow, K. J. 1951 *Social choice and individual values*. New York, NY: Wiley.

Bateson, M. 2002 Context-dependent foraging choices in risk-sensitive starlings. *Anim. Behav.* **64**, 251–260. (doi:10.1006/anbe.2002.3059)

Bateson, M., Healy, S. D. & Hurly, T. A. 2002 Irrational choices in hummingbird foraging behaviour. *Anim. Behav.* **63**, 587–596. (doi:10.1006/anbe.2001.1925)

Bateson, M., Healy, S. D. & Hurly, T. A. 2003 Context-dependent foraging decisions in rufous hummingbirds. *Proc. R. Soc. B* **270**, 1271–1276. (doi:10.1098/rspb.2003.2365)

Baum, W. M. 1974 On two types of deviation from matching laws: bias and under-matching. *J. Exp. Anal. Behav.* **22**, 231–242. (doi:10.1901/jeab.1974.22-231)

Belinsky, R., Gonzalez, F. & Stahl, J. 2004 Optimal behavior and concurrent variable interval schedules. *J. Math. Psychol.* **48**, 247–262. (doi:10.1016/j.jmp.2004.04.001)

Bernardo, A. E. & Welch, I. 2001 On the evolution of overconfidence and entrepreneurs. *J. Econ. Manag. Strategy* **10**, 301–330. (doi:10.1162/105864001316907964)

Blomquist, G. 1986 A utility maximization model of driver traffic safety behavior. *Accid. Anal. Prev.* **18**, 371–375. (doi:10.1016/0001-4575(86)90011-4)

Busemeyer, J. R. & Townsend, J. T. 1993 Decision field-theory—a dynamic cognitive approach to decision-making in an uncertain environment. *Psychol. Rev.* **100**, 432–459. (doi:10.1037/0033-295X.100.3.432)

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L. & Wakefield, J. C. 1998 Adaptations, exaptations, and spandrels. *Am. Psychol.* **53**, 533–548. (doi:10.1037/0003-066X.53.5.533)

Cervero, R. 2003 Road expansion, urban growth, and induced travel—a path analysis. *J. Am. Plan. Assoc.* **69**, 145–163.

Christen, U. & von Herrath, M. G. 2005 Infections and autoimmunity—good or bad? *J. Immunol.* **174**, 7481–7486.

Clark, C. W. 1994 Antipredator behavior and the asset-protection principle. *Behav. Ecol.* **5**, 159–170. (doi:10.1093/beheco/5.2.159)

Cooper, W. S. 2001 *The evolution of reason*. Cambridge, UK: Cambridge University Press.

Cuthill, I. C., Maddocks, S. A., Weall, C. V. & Jones, E. K. M. 2000 Body mass regulation in response to changes in feeding predictability and overnight energy expenditure. *Behav. Ecol.* **11**, 189–195. (doi:10.1093/beheco/11.2.189)

Dall, S. R. X., McNamara, J. M. & Cuthill, I. C. 1999 Interruptions to foraging and learning in a changing environment. *Anim. Behav.* **57**, 233–241. (doi:10.1006/anbe.1998.0944)

Darwin, C. 1839 *Journal of researches into the geology and natural history of the various countries visited by H. M. S. Beagle, under the command of Captain Fitzroy, R.N. from 1832 to 1836.* London, UK: Henry Colburn.

Day, T. & Proulx, S. R. 2004 A general theory for the evolutionary dynamics of virulence. *Am. Nat.* **163**, E40–E63. (doi:10.1086/382548)

Fantino, E. 1969 Choice and rate of reinforcement. *J. Exp. Anal. Behav.* **12**, 723–730. (doi:10.1901/jeab.1969.12-723)

Fantino, E. & Abarca, N. 1985 Choice, optimal foraging, and the delay-reduction hypothesis. *Behav. Brain Sci.* **8**, 315–329.

Fantino, E. & Davison, M. 1983 Choice—some quantitative relations. *J. Exp. Anal. Behav.* **40**, 1–13. (doi:10.1901/jeab.1983.40-1)

Fantino, E. & Dunn, R. 1983 The delay-reduction hypothesis—extension to three-alternative choice. *J. Exp. Psychol. Anim. Behav. Process.* **9**, 132–146. (doi:10.1037/0097-7403.9.2.132)

Fantino, E., Preston, R. A. & Dunn, R. 1993 Delay reduction—current status. *J. Exp. Anal. Behav.* **60**, 159–169. (doi:10.1901/jeab.1993.60-159)

Fawcett, T. W. & Johnstone, R. A. 2003 Optimal assessment of multiple cues. *Proc. R. Soc. B* **270**, 1637–1643. (doi:10.1098/rspb.2003.2328)

Ferster, C. B. & Skinner, B. F. 1957 *Schedules of reinforcement.* Englewood Cliffs, NJ: Prentice Hall.

Frank, S. A. 1996 Models of parasite virulence. *Q. Rev. Biol.* **71**, 37–78. (doi:10.1086/419267)

Fuqua, R. 1984 Comments on the applied relevance of the matching law. *J. Appl. Behav. Anal.* **17**, 381–386. (doi:10.1901/jaba.1984.17-381)

Goodwin, P. & Noland, R. B. 2003 Building new roads really does create extra traffic: a response to Prakash *et al. Appl. Econ.* **35**, 1451–1457. (doi:10.1080/0003684032000089872)

Gould, S. J. & Lewontin, R. C. 1979 The spandrels of San Marco and the Panglossian Paradigm—a critique of the adaptationist programme. *Proc. R. Soc. B* **205**, 581–598. (doi:10.1098/rspb.1979.0086)

Grafen, A. 1999 Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proc. R. Soc. B* **266**, 799–803. (doi:10.1098/rspb.1999.0708)

Haccou, P. & Iwasa, Y. 1995 Optimal mixed strategies in stochastic environments. *Theor. Popul. Biol.* **47**, 212–243. (doi:10.1006/tpbi.1995.1009)

Hammerstein, P. & Hagen, E. H. 2005 The second wave of evolutionary economics in biology. *Trends Ecol. Evol.* **20**, 604–609. (doi:10.1016/j.tree.2005.07.012)

Hampton, S. J. 2004 Domain mismatches, scruffy engineering, exaptations and spandrels. *Theory Psychol.* **14**, 147–166. (doi:10.1177/0959354304042014)

Harley, C. B. 1981 Learning the evolutionarily stable strategy. *J. Theor. Biol.* **89**, 611–633. (doi:10.1016/0022-5193(81)90032-1)

Herrnstein, R. J. 1961 Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267–272. (doi:10.1901/jeab.1961.4-267)

Herrnstein, R. J. 1970 On the law of effect. *J. Exp. Anal. Behav.* **13**, 244–266. (doi:10.1901/jeab.1970.13-243)

Hey, J. D. 2002 Experimental economics and the theory of decision making under risk and uncertainty. *Geneva Pap. Risk Ins.* **27**, 5–21. (doi:10.1023/A:1020634723329)

Heyman, G. M. & Herrnstein, R. J. 1986 More on concurrent interval-ratio schedules—a replication and review. *J. Exp. Anal. Behav.* **46**, 331–351. (doi:10.1901/jeab.1986.46-331)

Heyman, G. M. & Luce, R. D. 1979 Operant matching is not a logical consequence of maximizing reinforcement rate. *Anim. Learn. Behav.* **7**, 133–140.

Heyman, G. M. & Tanz, L. 1995 How to teach a pigeon to maximize overall reinforcement rate. *J. Exp. Anal. Behav.* **64**, 277–297. (doi:10.1901/jeab.1995.64-277)

Houston, A. I. 1983 Optimality theory and matching. *Behav. Anal. Lett.* **3**, 1–15.

Houston, A. I. 1986a The matching law applies to wagtails' foraging in the wild. *J. Exp. Anal. Behav.* **45**, 15–18. (doi:10.1901/jeab.1986.45-15)

Houston, A. I. 1986b The optimal flight velocity for a bird exploiting patches of food. *J. Theor. Biol.* **119**, 345–362. (doi:10.1016/S0022-5193(86)80146-1)

Houston, A. I. 1987 The control of foraging decisions. In *Quantitative analysis of behaviour* (eds M. L. Commons, A. Kacelnik & S. J. Shettleworth), pp. 41–61. Hillsdale, NJ: Lawrence Erlbaum.

Houston, A. I. 1991 Violations of stochastic transitivity on concurrent chains: implications for theories and choice. *J. Exp. Anal. Behav.* **55**, 323–335. (doi:10.1901/jeab.1991.55-323)

Houston, A. I. 1997 Natural selection and context-dependent values. *Proc. R. Soc. B* **264**, 1539–1541. (doi:10.1098/rspb.1997.0213)

Houston, A. I. & Carbone, C. 1992 The optimal allocation of time during the diving cycle. *Behav. Ecol.* **3**, 255–265. (doi:10.1093/beheco/3.3.255)

Houston, A. I. & McNamara, J. 1981 How to maximize reward rate on 2 variable-interval paradigms. *J. Exp. Anal. Behav.* **35**, 367–396. (doi:10.1901/jeab.1981.35-367)

Houston, A. I. & McNamara, J. M. 1988 A framework for the functional analysis of behavior. *Behav. Brain Sci.* **11**, 117–154.

Houston, A. I. & McNamara, J. M. 1989 The value of food: effects of open and closed economies. *Anim. Behav.* **37**, 546–562. (doi:10.1016/0003-3472(89)90034-1)

Houston, A. I. & McNamara, J. M. 1999 *Models of adaptive behaviour: an approach based on state.* Cambridge, UK: Cambridge University Press.

Houston, A. I. & Staddon, J. E. R. 1981 Optimality principles and behavior: it's all for the best. *Behav. Brain Sci.* **4**, 395–396.

Houston, A. I. & Sumida, B. H. 1987 Learning rules, matching and frequency-dependence. *J. Theor. Biol.* **126**, 289–308. (doi:10.1016/S0022-5193(87)80236-9)

Houston, A. I., McNamara, J. M. & Hutchinson, J. M. C. 1993 General results concerning the trade-off between gaining energy and avoiding predation. *Phil. Trans. R. Soc. B* **341**, 375–397. (doi:10.1098/rstb.1993.0123)

Houston, A. I., McNamara, J. M., Heron, J. E. & Barta, Z. 2003 The effect of foraging parameters on the probability that a dive is successful. *Proc. R. Soc. B* **270**, 2451–2455. (doi:10.1098/rspb.2003.2540)

Houston, A. I., Steer, M. D., Killeen, P. R. & Thompson, W. A. 2005 Predicting violations of transitivity when choices involve fixed or variable delays to food. In *Modelling natural action selection: an international workshop, Edinburgh* (eds J. J. Bryson, T. J. Prescott & A. K. Seth), pp. 9–15. Edinburgh, UK: AISB Press.

Houston, A. I., McNamara, J. M. & Steer, M. D. 2007 Violations of transitivity under fitness maximization. *Biol. Lett.* **3**, 365–367. (doi:10.1098/rsbl.2007.0111)

Hurly, T. A. & Oseen, M. D. 1999 Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. *Anim. Behav.* **58**, 59–66. (doi:10.1006/anbe.1999.1130)

Hutchinson, J. M. C. & Gigerenzer, G. 2005 Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behav. Processes* **69**, 97–124. (doi:10.1016/j.beproc.2005.02.019)

Iwasa, Y., Higashi, M. & Yamamura, N. 1981 Prey distribution as a factor determining the choice of optimal foraging strategy. *Am. Nat.* **117**, 710–723. (doi:10.1086/283754)

Janssen, W. H. & Tenkink, E. 1988 Considerations on speed selection and risk homeostasis in driving. *Accid. Anal. Prev.* **20**, 137–142. (doi:10.1016/0001-4575(88)90030-9)

Jokela, J., Schmid-Hempel, P. & Rigby, M. C. 2000 Dr. Pangloss restrained by the Red Queen—steps towards a unified defence theory. *Oikos* **89**, 267–274. (doi:10.1034/j.1600-0706.2000.890207.x)

Kacelnik, A. 2006 Meanings of rationality. In *Rational animals?* (eds S. Hurley & M. Nudds). Oxford, UK: Oxford University Press.

Kacelnik, A. & Bateson, M. 1996 Risky theories—the effects of variance on foraging decisions. *Am. Zool.* **36**, 402–434.

Kahneman, D. 2003 Maps of bounded rationality: psychology for behavioral economics. *Am. Econ. Rev.* **93**, 1449–1475. (doi:10.1257/000282803322655392)

Keeler, T. E. 1994 Highway safety, economic-behavior, and driving environment. *Am. Econ. Rev.* **84**, 684–693.

Kramer, D. L. 1988 The behavioral ecology of air breathing by aquatic animals. *Can. J. Zool.* **66**, 89–94.

Lima, S. L. 1998 Nonlethal effects in the ecology of predator–prey interactions—what are the ecological effects of anti-predator decision-making? *Bioscience* **48**, 25–34. (doi:10.2307/1313225)

Luce, R. D. 1959 *Individual choice behaviour: a theoretical analysis.* New York, NY: Wiley.

Luce, R. D. 1977 The choice axiom after twenty years. *J. Math. Psychol.* **15**, 215–233. (doi:10.1016/0022-2496(77)90032-3)

Luce, R. D. & Raiffa, H. 1957 *Games and decisions.* New York, NY: Wiley.

Luce, R. D. & Suppes, P. 1965 Preference, utility, and subjective probability. In *Handbook of psychology III* (eds R. D. Luce, R. R. Bush & E. Galanter), pp. 249–410. New York, NY: Wiley.

Manktelow, K. I. & Over, D. E. 1993 *Rationality: psychological and philosophical perspectives*. London, UK; New York, NY: Routledge.

March, J. G. 1996 Learning to be risk averse. *Psychol. Rev.* **103**, 309–319. (doi:10.1037/0033-295X.103.2.309)

Mazur, J. E. 2000 Two- versus three-alternative concurrent-chain schedules: a test of three models. *J. Exp. Psychol. Anim. Behav. Process.* **26**, 286–293. (doi:10.1037/0097-7403.26.3.286)

McDowell, J. J. 1981 On the validity and utility of Herrnstein's hyperbola in applied behavior analysis. In *Quantification of steady-state operant behavior* (eds C. M. Bradshaw, E. Szabadi & C. F. Lowe). Amsterdam, The Netherlands: Elsevier.

McDowell, J. J. 1989 Two modern developments in matching theory. *Behav. Anal.* **12**, 153–166.

McNamara, J. M. 1982 Optimal patch use in a stochastic environment. *Theor. Popul. Biol.* **21**, 269–288. (doi:10.1016/0040-5809(82)90018-1)

McNamara, J. M. 1990 The policy which maximizes long-term survival of an animal faced with the risks of starvation and predation. *Adv. Appl. Probab.* **22**, 295–308. (doi:10.2307/1427537)

McNamara, J. M. 1995 Implicit frequency-dependence and kin selection in fluctuating environments. *Evol. Ecol.* **9**, 185–203. (doi:10.1007/BF01237756)

McNamara, J. M. 1996 Risk-prone behaviour under rules which have evolved in a changing environment. *Am. Zool.* **36**, 484–495.

McNamara, J. M. & Buchanan, K. L. 2005 Stress, resource allocation, and mortality. *Behav. Ecol. Sociobiol.* **16**, 1008–1017.

McNamara, J. M. & Houston, A. I. 1980 The application of statistical decision-theory to animal behaviour. *J. Theor. Biol.* **85**, 673–690. (doi:10.1016/0022-5193(80)90265-9)

McNamara, J. M. & Houston, A. I. 1985 Optimal foraging and learning. *J. Theor. Biol.* **117**, 231–249. (doi:10.1016/S0022-5193(85)80219-8)

McNamara, J. M. & Houston, A. I. 1986 The common currency for behavioral decisions. *Am. Nat.* **127**, 358–378. (doi:10.1086/284489)

McNamara, J. M. & Houston, A. I. 1987a Partial preferences and foraging. *Anim. Behav.* **35**, 1084–1099. (doi:10.1016/S0003-3472(87)80166-5)

McNamara, J. M. & Houston, A. I. 1987b Memory and the efficient use of information. *J. Theor. Biol.* **125**, 385–395. (doi:10.1016/S0022-5193(87)80209-6)

McNamara, J. M. & Houston, A. I. 1987c Foraging in patches—there's more to life than the marginal value theorem. In *Quantitative analyses of behavior* (eds M. L. Commons, A. Kacelnik & S. J. Shettleworth), pp. 23–39. Hillsdale, NJ: Lawrence Erlbaum Associates.

McNamara, J. M. & Houston, A. I. 1987d Starvation and predation as factors limiting population size. *Ecology* **68**, 1515–1519. (doi:10.2307/1939235)

McNamara, J. M. & Houston, A. I. 1990a The value of fat reserves and the trade-off between starvation and predation. *Acta Biotheor.* **38**, 37–61. (doi:10.1007/BF00047272)

McNamara, J. M. & Houston, A. I. 1990b Starvation and predation in a patchy environment. In *Living in a patchy environment* (eds I. R. Swingland & B. Shorrocks), pp. 23–43. Oxford, UK: Oxford University Press.

McNamara, J. M. & Houston, A. I. 1992 Risk-sensitive foraging—a review of the theory. *Bull. Math. Biol.* **54**, 355–378.

McNamara, J. M. & Houston, A. I. 1994 The effect of a change in foraging options on intake rate and predation rate. *Am. Nat.* **144**, 978–1000. (doi:10.1086/285721)

McNamara, J. M., Houston, A. I. & Collins, E. J. 2001 Optimality models in behavioural biology. *SIAM Rev.* **43**, 413–466. (doi:10.1137/S0036144500385263)

McNamara, J. M., Barta, Z., Houston, A. I. & Race, P. 2005 A theoretical investigation of the effect of predators on foraging behaviour and energy reserves. *Proc. R. Soc. B* **272**, 929–934. (doi:10.1098/rspb.2004.3037)

Medley, G. F. 2002 The epidemiological consequences of optimisation of the individual host immune response. *Parasitology* **125**, S61–S70.

Moser, P. K. 1990 *Rationality in action: contemporary approaches*. Cambridge, UK: Cambridge University Press.

Myerson, J. & Hale, S. 1984 Practical implications of the matching law. *J. Appl. Behav. Anal.* **17**, 367–380. (doi:10.1901/jaba.1984.17-367)

Navarick, D. J. & Fantino, E. 1972 Transitivity as a property of choice. *J. Exp. Anal. Behav.* **18**, 389–401. (doi:10.1901/jeab.1972.18-389)

Navarick, D. J. & Fantino, E. 1974 Stochastic transitivity and unidimensional behavior theories. *Psychol. Rev.* **81**, 426–441. (doi:10.1037/h0036953)

Navarick, D. J. & Fantino, E. 1975 Stochastic transitivity and unidimensional control of choice. *Learn. Motiv.* **6**, 179–201. (doi:10.1016/0023-9690(75)90021-1)

Noland, R. B. & Lem, L. L. 2002 A review of the evidence for induced travel and changes in transportation and environmental policy in the US and the UK. *Transp. Res. Part D Transp. Environ.* **7**, 1–26. (doi:10.1016/S1361-9209(01)00009-8)

Norberg, R. Å. 1981 Optimal flight speed in birds when feeding young. *J. Anim. Ecol.* **50**, 473–477. (doi:10.2307/4068)

Oster, G. F. & Wilson, E. O. 1978 *Caste and ecology in the social insects*. Princeton, NJ: Princeton University Press.

Peltzman, S. 1975 Effects of automobile safety regulation. *J. Polit. Econ.* **83**, 677–725. (doi:10.1086/260352)

Peterson, S. P. & Hoffer, G. E. 1994 The impact of airbag adoption on relative personal-injury and absolute collision insurance claims. *J. Consum. Res.* **20**, 657–662. (doi:10.1086/209377)

Pierce, W. D. & Epling, W. F. 1995 The applied importance of research on the matching law. *J. Appl. Behav. Anal.* **28**, 237–241. (doi:10.1901/jaba.1995.28-237)

Rieskamp, J., Busemeyer, J. R. & Mellers, B. A. 2006 Extending the bounds of rationality: evidence and theories of preferential choice. *J. Econ. Lit.* **44**, 631–661. (doi:10.1257/jel.44.3.631)

Robson, A. J. 2002 Evolution and human nature. *J. Econ. Perspect.* **16**, 89–106.

Robson, A. J. 2003 The evolution of rationality and the Red Queen. *J. Econ. Theory* **111**, 1–22. (doi:10.1016/S0022-0531(03)00084-X)

Romagnani, S. 2004 The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? *Immunology* **112**, 352–363. (doi:10.1111/j.1365-2567.2004.01925.x)

Sasaki, A. & Ellner, S. 1995 The evolutionarily stable phenotype distribution in a random environment. *Evolution* **49**, 337–350. (doi:10.2307/2410344)

Savastano, H. I. & Fantino, E. 1994 Human choice in concurrent ratio-interval schedules of reinforcement. *J. Exp. Anal. Behav.* **61**, 453–463. (doi:10.1901/jeab.1994.61-453)

Schuck-Paim, C., Pompilio, L. & Kacelnik, A. 2004 State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS. Biol.* **2**, 2305–2315. (doi:10.1371/journal.pbio.0020402)

Seth, A. K. 1999 Evolving behavioural choice: an investigation into Herrnstein's matching law. In *Advances in artificial life, proceedings* (eds D. Floreano *et al.*), pp. 225–235. Berlin, Germany: Springer Verlag.

Seth, A. K. 2001 Modeling group foraging: individual suboptimality, interference, and a kind of matching. *Adapt. Behav.* **9**, 67–89.

Seth, A. 2007 The ecology of action selection: insights from artificial life. *Phil. Trans. R. Soc. B* **362**, 1545–1558. (doi:10.1098/rstb.2007.2052)

Shafir, S. 1994 Intransitivity of preferences in honeybees—support for comparative-evaluation of foraging options. *Anim. Behav.* **48**, 55–67. (doi:10.1006/anbe.1994.1211)

Shafir, S., Waite, T. A. & Smith, B. H. 2002 Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behav. Ecol. Sociobiol.* **51**, 180–187. (doi:10.1007/s00265-001-0420-8)

Shettleworth, S. J. 1985 Foraging, memory, and constraints on learning. *Ann. N. Y. Acad. Sci.* **443**, 216–226. (doi:10.1111/j.1749-6632.1985.tb27075.x)

Shettleworth, S. J. 1989 Animals foraging in the lab—problems and promises. *J. Exp. Psychol. Anim. Behav. Process.* **15**, 81–87. (doi:10.1037/0097-7403.15.1.81)

Shettleworth, S. J. & Jordan, V. 1986 Rats prefer handling food to waiting for it. *Anim. Behav.* **34**, 925–927. (doi:10.1016/S0003-3472(86)80079-3)

Shudo, E. & Iwasa, Y. 2001 Inducible defense against pathogens and parasites: optimal choice among multiple options. *J. Theor. Biol.* **209**, 233–247. (doi:10.1006/jtbi.2000.2259)

Shudo, E. & Iwasa, Y. 2004 Dynamic optimization of host defense, immune memory, and post-infection pathogen levels in mammals. *J. Theor. Biol.* **228**, 17–29. (doi:10.1016/j.jtbi.2003.12.001)

Silberberg, A., Thomas, J. R. & Berendzen, N. 1991 Human choice on concurrent variable-interval variable-ratio schedules. *J. Exp. Anal. Behav.* **56**, 575–584. (doi:10.1901/jeab.1991.56-575)

Simon, H. A. 1976 From substantive to procedural rationality. In *Method and appraisal in economics* (ed. J. P. Latsis). Cambridge, UK: Cambridge University Press.

Smith, V. L. 2003 Constructivist and ecological rationality in economics. *Am. Econ. Rev.* **93**, 465–508.

Soltani, A. & Wang, X. J. 2006 A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* **26**, 3731–3744. (doi:10.1523/JNEUROSCI.5159-05.2006)

Spiga, R., Maxwell, R. S., Meisch, R. A. & Grabowski, J. 2005 Human methadone self-administration and the generalized matching law. *Psychol. Rec.* **55**, 525–538.

Squires, N. & Fantino, E. 1971 A model for choice in simple concurrent and concurrent-chains schedules. *J. Exp. Anal. Behav.* **15**, 27–38. (doi:10.1901/jeab.1971.15-27)

Steer, M. D. & Cuthill, I. C. 2003 Irrationality, suboptimality, and the evolutionary context. *Behav. Brain Sci.* **26**, 176–177. (doi:10.1017/S0140525X03460056)

Stephens, D. W. & Krebs, J. R. 1986 *Foraging theory.* Princeton, NJ: Princeton University Press.

Stephens, D. W., Kerr, B. & Fernandez-Juricic, E. 2004 Impulsiveness without discounting: the ecological rationality hypothesis. *Proc. R. Soc. B* **271**, 2459–2465. (doi:10.1098/rspb.2004.2871)

Strachan, D. P. 1989 Hay-fever, hygiene, and household size. *Br. Med. J.* **299**, 1259–1260.

Sugrue, L. P., Corrado, G. S. & Newsome, W. T. 2004 Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787. (doi:10.1126/science.1094765)

Sutherland, S. 1994 *Irrationality: the enemy within.* London, UK: Penguin Books.

Tenner, E. 1996 *Why things bite back: new technology and the revenge effect.* London, UK: Fourth Estate.

Todd, P. M. & Gigerenzer, G. 2000 How can we open up the adaptive toolbox? *Behav. Brain Sci.* **23**, 767–780.

Tversky, A. & Kahneman, D. 1981 The framing of decisions and the psychology of choice. *Science* **211**, 453–458. (doi:10.1126/science.7455683)

Tversky, A. & Simonson, I. 1993 Context-dependent preferences. *Manage. Sci.* **39**, 1179–1189.

van Boven, M. & Weissing, F. J. 2004 The evolutionary economics of immunity. *Am. Nat.* **163**, 277–294. (doi:10.1086/381407)

VanDam, N. M., DeJong, T. J., Iwasa, Y. & Kubo, T. 1996 Optimal distribution of defences: are plants smart investors? *Funct. Ecol.* **10**, 128–136. (doi:10.2307/2390271)

Waite, T. A. 2001*a* Background context and decision making in hoarding gray jays. *Behav. Ecol.* **12**, 318–324. (doi:10.1093/beheco/12.3.318)

Waite, T. A. 2001*b* Intransitive preferences in hoarding gray jays (*Perisoreus canadensis*). *Behav. Ecol. Sociobiol.* **50**, 116–121. (doi:10.1007/s002650100346)

Waite, T. A. & Passino, K. M. 2006 Paradoxical preferences when options are identical. *Behav. Ecol. Sociobiol.* **59**, 777–785. (doi:10.1007/s00265-005-0123-7)

Welton, N. J. & Houston, A. I. 2001 A theoretical investigation into the direct and indirect effects of state on the risk of predation. *J. Theor. Biol.* **123**, 275–297. (doi:10.1006/jtbi.2001.2419)

Wilson, B. 1974 *Rationality.* Oxford, UK: Basil Blackwell.

Yazdanbakhsh, M., Kremsner, P. G. & van Ree, R. 2002 Immunology—allergy, parasites, and the hygiene hypothesis. *Science* **296**, 490–494. (doi:10.1126/science.296.5567.490)