

Invited Review

# Exposing the behavioral gambit: the evolution of learning and decision rules

Tim W. Fawcett,<sup>a,b,c</sup> Steven Hamblin,<sup>b</sup> and Luc-Alain Giraldeau<sup>b</sup>

<sup>a</sup>GRIP, Université de Montréal, 3050 Édouard-Montpetit, Montréal (QC) H3T 1J7, Canada

<sup>b</sup>Département des Sciences Biologiques, Université du Québec à Montréal, CP 8888, Succ. Centre-ville, Montréal (QC) H3C 3P8, Canada, and <sup>c</sup>Modelling Animal Decisions (MAD) Group, School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

Behavioral ecologists have long been comfortable assuming that genetic architecture does not constrain which phenotypes can evolve (the “phenotypic gambit”). For flexible behavioral traits, however, solutions to adaptive problems are reached not only by genetic evolution but also by behavioral changes within an individual’s lifetime, via psychological mechanisms such as learning. Standard optimality approaches ignore these mechanisms, implicitly assuming that they do not constrain the expression of adaptive behavior. This assumption, which we dub the behavioral gambit, is sometimes wrong: evolved psychological mechanisms can prevent animals from behaving optimally in specific situations. To understand the functional basis of behavior, we would do better by considering the underlying mechanisms, rather than the behavioral outcomes they produce, as the target of selection. This change of focus yields new, testable predictions about evolutionary equilibria, the development of behavior, and the properties of cognitive systems. Studies on the evolution of learning rules hint at the potential insights to be gained, but such mechanism-based approaches are underexploited. We highlight three future research priorities: (1) systematic theoretical analysis of the evolutionary properties of learning rules; (2) detailed empirical study of how animals learn in nonforaging contexts; and (3) analysis of individual differences in learning rules and their associated fitness consequences. *Key words:* behavioral flexibility, behaviorally stable strategy, behavioral plasticity, game theory, learning rules, optimality. [*Behav Ecol*]

## INTRODUCTION

Perhaps the greatest success of behavioral ecology has been its ability to make detailed predictions of animal behavior from an economic consideration of the costs and benefits associated with alternative actions (Parker and Maynard Smith 1990; Krebs and Davies 1997). Underpinning this approach is the so-called phenotypic gambit—the assumption that genetic architecture does not constrain which phenotypes can evolve (Grafen 1984). There is continuing debate about the validity of this gambit (Hadfield et al. 2007; Bull and Wang 2010), but by and large, it is regarded as a useful approximation (Cheverud 1988; Roff 1995; Réale and Festa-Bianchet 2000). Although genetics may make a difference in short-term evolution, they can perhaps be safely neglected in the long term (Weissing 1996).

Showcasing the power of the phenotypic gambit are the polymorphisms of marine isopods (*Paracerceis sculpta*; Shuster and Wade 1991), scale-eating cichlids (*Perissodus microlepis*; Hori 1993), and side-blotched lizards (*Uta stansburiana*; Sinervo and Lively 1996), in which the frequencies of alternative behavioral strategies can be predicted from phenotypic measurements of frequency-dependent selection. In all these systems, the behavioral strategies are more or less fixed from

birth and are accompanied by marked physiological and morphological differentiation. Behavioral changes occur from one generation to the next through selection and inheritance of alternative strategies, resulting in genetic evolution.

Most behaviors of interest, however, are not fixed but malleable. In these cases, adaptation occurs not only through genetic evolution but also through changes that take place within an individual’s lifetime, via psychological mechanisms such as learning. Despite this important distinction, functional analyses of behavior typically treat flexible and fixed behaviors in the same way. The methods used to analyze moment-to-moment behavioral decisions, such as where to forage, whom to interact with, or whether to explore or exploit the environment, are the same as those used to analyze life-history decisions over clutch size, sex allocation, and the timing of maturity. This approach invokes an additional, unstated assumption: that the psychological mechanisms underlying flexible decision making do not constrain the expression of adaptive behavior and allow animals to reach the optimal solution to a given problem. We refer to this extra assumption as the behavioral gambit (Giraldeau and Dubois 2008; see also Puts 2010). Put simply, “sophisticated behavioral adaptations of animals are thought to reflect the calculation power of the evolutionary process, rather than cognitive skills of the individual brain” (Hammerstein 1998, p. 5). Here, we argue that the way that evolution has shaped cognitive systems may in fact have a crucial impact on behavior.

By focusing on expressed behavior and neglecting the underlying mechanism, behavioral ecologists unwittingly adopt the behavioral gambit, extending the phenotypic gambit beyond its accepted remit. For example, when predicting the amount of food that a central-place forager should collect

Address correspondence to T.W. Fawcett. E-mail: [tim.fawcett@can-tab.net](mailto:tim.fawcett@can-tab.net).

S. Hamblin is now at the School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia.

Received 8 February 2011; revised 14 July 2011; accepted 17 April 2012.

before returning to its nest, researchers typically omit the details of how the animal perceives its environment, how it values food items, how it remembers its past foraging experiences, and how it devalues the future, assuming that all of these have evolved to provide the animal with the optimal behavioral solution. Mechanisms have not been ignored, but by and large, they are studied in parallel and almost never integrated with a functional perspective on animal behavior (cf. Brunner et al. 1996; Stephens 2002). This can be a useful simplification, but, as we argue in this article, it has its limitations. We believe it is critical that researchers recognize when they rely on the behavioral gambit for two key reasons. First, the behavioral gambit sometimes fails: evolved psychological mechanisms can prevent animals from behaving optimally in specific situations (see e.g. Herrnstein 1990; Bateson and Healy 2005; Henly et al. 2008). To understand why the gambit could be wrong, we need to consider the adaptive value (i.e. fitness-affecting properties) of the underlying mechanisms, not the behaviors they produce (Stevens 2008). As we will see, predictions of behavior based on evolved mechanisms do not always coincide with standard predictions, revealing the shortcomings of the gambit. Second, by questioning the behavioral gambit and shifting our focus to psychological mechanisms as the phenotype under selection, we can obtain new insights about adaptive behavior. We ask what kinds of psychological mechanisms selection should favor, focusing particularly on the evolution of learning rules and providing a critical appraisal of previous research on this topic. Finally, we offer recommendations for future research on flexible behavior.

### IS THE BEHAVIORAL GAMBIT JUSTIFIED?

Recognizing our reliance on the behavioral gambit is crucial because sometimes it is wrong. Experimental psychologists have analyzed in detail how laboratory animals make decisions when faced with alternative options, and quite often the observed behavior does not fit predictions from a standard optimality approach. Below we describe three illustrative examples of such apparently maladaptive behavior.

#### Animals are impulsive

Consider an animal given a choice between a smaller amount of food immediately and a much larger amount of food after a brief time delay, the so-called self-control paradigm (Stephens et al. 2004). When the larger, later option gives a higher overall rate of intake, optimal foraging theory predicts that the animal should prefer this option. In reality, however, animals are commonly found to be impulsive; that is, they prefer the immediate reward (reviewed in Henly et al. 2008). One popular explanation for this puzzling behavior is that delayed rewards are devalued by the risk that foraging will be interrupted, thereby preventing the animal from collecting the reward (Kagel et al. 1986), but studies of temporal discounting do not support this interpretation (Mazur 2000; Henly et al. 2008).

#### Animals do not maximize their rewards

When the reward rates associated with different behaviors are unequal, animals tend to adjust their use of those behaviors accordingly; behaviors that are rewarded at a high rate are performed more frequently than those that are less profitable (a phenomenon known as melioration; Herrnstein and Vaughan 1980). The resulting stable pattern of allocation is well described by Herrnstein's (1961) "matching law," which

states that the proportion of times that an option is chosen is equal to the proportion of reinforcements that the option delivers. Since Herrnstein's classic study, the generality of his law has been confirmed by experiments on a range of species (Davison and McCarthy 1988; Staddon and Cerutti 2003). Strikingly, animals appear to show matching even when alternative patterns of behavior would yield a higher rate of gain (Herrnstein and Heyman 1979; Houston and McNamara 1981; Heyman and Herrnstein 1986; Herrnstein 1990). This is puzzling from an optimality perspective, under which (all else being equal) we expect animals to maximize their rate of gain.

#### Animals do not value alternative options in a rational way

Rationality is the cornerstone of economic models (Kacelnik 2006; Houston et al. 2007), yet animals often do not value alternative options in a rational way. Sometimes valuation is state dependent: it is affected by the condition the animal was in when it was previously rewarded for that behavior. For example, European starlings (*Sturnus vulgaris*; Marsh et al. 2004), desert locusts (*Schistocerca gregaria*; Pompilio et al. 2006), and banded tetras (*Astyanax fasciatus*; Aw et al. 2009) prefer an option which previously gave them food when they were hungry over one that gave them the same amount of food when they were satiated, whereas laboratory rats (*Rattus norvegicus*) prefer a flavor more strongly if they originally tasted it under low rather than high hunger (Capaldi et al. 1983). The choices animals make are also influenced by irrelevant or "decoy" options (Bateson and Healy 2005). For example, in honeybees (*Apis mellifera*; Shafir et al. 2002), grey jays (*Perisoreus canadensis*; Shafir et al. 2002), European starlings (Bateson 2002), and rufous hummingbirds (*Selasphorus rufus*; Bateson et al. 2003), the preference for one of two foraging options increases when a third, relatively unattractive option is added to the choice set. This violates principles of rational choice and implies that valuation is not absolute, but comparative (though see Schuck-Paim et al. 2004 for a different interpretation). From a standard optimality perspective on behavior, such state- and context-dependent valuation appears not to be adaptive. Thus, in contradiction to the behavioral gambit, the underlying psychological mechanisms have not evolved to support optimal outcomes.

### EVOLVED MECHANISMS AS ADAPTATION AND CONSTRAINT

Why is behavior not adaptive in these cases? We propose that this is because the behavior patterns expressed in specific situations are not themselves directly under selection but are the product of evolved psychological mechanisms underlying behavioral flexibility across a wide range of situations (Stevens 2008). Natural environments are so complex, dynamic, and unpredictable that natural selection cannot possibly furnish an animal with an appropriate, specific behavior pattern for every conceivable situation it might encounter (McNamara and Houston 2009). Instead, we should expect animals to have evolved a set of psychological mechanisms which enable them to perform well on average across a range of different circumstances (Gigerenzer et al. 1999; Hutchinson and Gigerenzer 2005). These mechanisms encompass fixed rules for responding to current stimuli, subject to sensory biases, and learning rules for how to adjust behavior in response to past experiences.

The evolved psychological mechanisms of a given species will be tuned to its ecology, reflecting the particular kinds of situations the animal is likely to confront in its

natural environment. Many situations will never have been encountered before, but the animal may rely on the same mechanisms for tackling problems with a similar statistical structure (Todd and Gigerenzer 2007; Wilke and Barrett 2009). In certain specific situations—including the contrived laboratory situations common in experimental psychology—the behavior displayed may not maximize rate of energetic gain, or whichever other currency is assumed to influence fitness. To understand why that behavior has evolved, we have to consider the adaptive value of the psychological mechanism which controls it (Stevens 2008; McNamara and Houston 2009), in the kinds of environments the animal would normally encounter (McNamara and Houston 1980; Houston and McNamara 1989; Houston et al. 2007; Todd and Gigerenzer 2007; Houston 2009).

Most behavioral ecologists do not directly address psychological mechanisms when considering the adaptive value of behavior, but work led by Alex Kacelnik (e.g. Kacelnik and Todd 1992; Todd and Kacelnik 1993; Brunner et al. 1996; Shapiro et al. 2008) and Dave Stephens (e.g. Stephens and Anderson 2001; Stephens 2002; Stephens et al. 2004) provides a prominent exception. For example, Brunner et al. (1996) incorporated memory constraints into an optimality model of patch-leaving decisions in starlings, inspired by empirical research on how animals estimate time intervals (Gibbon 1977). This model predicted the empirical data better than one in which the animals were assumed to have complete memory of all the time intervals between the food items found in a patch (Brunner et al. 1996). In a similar vein, a model by Stephens (2002) assumed that a given absolute difference between two stimuli is easier to discern when those stimuli are of low intensity, a well-known empirical result known as Weber's law (Gibbon 1977). He used this to show that impulsive decisions could, under certain conditions, be adaptive (Stephens 2002).

This kind of approach highlights the potential benefits of incorporating psychological mechanisms into standard optimality models, but it is rarely adopted. Furthermore, the emphasis in these examples is firmly on psychological mechanisms as a constraint on optimization. A particular mechanistic constraint is assumed on the basis of empirical findings, and then the aim is to find the optimal strategy operating within those constraints. Here, we contend that the mechanisms underlying behavioral flexibility should be viewed not only as a constraint but also as a possible adaptation. It was clear from the preceding section that psychological mechanisms may sometimes constrain an animal's behavior in such a way that it performs suboptimally within particular situations. But a mechanism that consistently produces suboptimal

behavior in common situations will be selected against, particularly if the deviations from optimal behavior are very costly (Houston 1987). We therefore expect that the set of psychological mechanisms an animal possesses will, overall, be well adapted to its natural environment, generating behavior which is close to optimal in the kinds of situations the animal is most likely to face.

## CHECKING OUR PREDICTIONS OF BEHAVIOR

In adopting the behavioral gambit, mechanistic constraints are put aside and are assumed not to affect the behavioral outcome. This is a convenient assumption because behavior is observable, whereas the underlying psychology is hidden, but as we have seen above it does not always work. There is an interesting parallel here with the economists' metaphor of *Homo economicus*, the idealized being who makes perfectly rational decisions, free from the complications of emotions or feelings (Persky 1995). As a predictive model of human decision making, this has been criticized, partly because emotions are known to play a pivotal role in many of the decisions made by real people (Loewenstein 2000) and may sometimes cause them to behave irrationally (Henrich et al. 2001; Shiv et al. 2005). For example, Shiv et al. (2005) found that patients with damage to the brain regions involved in the processing of emotions performed better on a monetary investment task than both normal participants and patients with damage to other brain regions. This suggests that emotions can sometimes "interfere" with rational decision making. In line with what we have argued above, a way to understand such instances of irrational behavior would be to consider whether the underlying psychological mechanisms are adaptive in a broader context.

Has our unquestioning acceptance of the behavioral gambit led us to incorrect predictions about animal behavior? To find this out, we need to revisit standard predictions and check whether evolved psychological mechanisms would lead to the same solution. This requires that we focus on the mechanisms, rather than the behavioral outcomes they produce, as the target of selection. This perspective turns out to be highly instructive and can lead to new insights about adaptive behavior.

## Reassessing evolutionary equilibria

In game-theoretical situations, where pay-offs are influenced by the actions of others, natural selection is expected to lead to an equilibrium state at which no alternative types can spread

**Table 1**  
Definitions of some key equilibrium concepts in behavioral ecology

Term	Definition	Reference
Stable equilibrium frequency (SEF)	Within a population, a phenotype or stable mixture of phenotypes for which no individual would increase its fitness by changing to a different phenotype. No distinction is made between whether the phenotypes are fixed (stable from birth) or flexible (adjustable within an individual's lifetime).	Mottley and Giraldeau (2000)
Evolutionarily stable strategy (ESS)	An SEF for which individual phenotypes are assumed (implicitly or explicitly) to be genetically fixed and held in equilibrium by frequency-dependent selection.	Maynard Smith (1982)
Behaviorally stable strategy (BSS)	An SEF for which individual phenotypes are flexible and can change during an individual's lifetime. The equilibrium may be reached via learning from frequency-dependent pay-offs.	Giraldeau and Dubois (2008)
Evolutionarily stable (ES) learning rule	A rule for adjusting behavior in response to past experiences which, when adopted by all individuals in a population, cannot be outcompeted by any alternative rule.	Harley (1981)



within the population, referred to as an evolutionarily stable strategy (ESS; [Maynard Smith 1982](#)) or stable equilibrium frequency (SEF; [Mottley and Giraldeau 2000](#); see [Table 1](#)). To predict such equilibria, alternative behavioral types are treated as though they are genetically fixed, with their proportions changing from one generation to the next through frequency-dependent selection. This basic methodology has been used to analyze behavior in a range of interactive contexts, from fighting to mating to cooperating.

However, in most of the behavioral games studied, a change in the frequencies of alternative strategies in the population is achieved not by selection weeding out less adapted forms but by individuals flexibly changing their use of those strategies ([Stephens and Clements 1998](#)). In such situations, an equilibrium is reached when no individual can benefit by unilaterally changing its behavior. This Nash equilibrium is analogous to an ESS but is more properly termed a behaviorally stable strategy (BSS) ([Giraldeau and Dubois 2008](#); [Dobler and Kölliker 2009](#); [Morand-Ferron and Giraldeau 2010](#); also developmentally stable strategy [DSS], [Dawkins 1976](#)) because it is reached through behavioral, rather than evolutionary, changes. In [Table 1](#), we distinguish between the different equilibrium concepts used in behavioral ecology. The behavioral gambit implies that an equilibrium state will be the same regardless of whether it is reached over evolutionary time (ESS) or through individuals changing their behavior (BSS). But is this supported? To answer this question, we must consider the rules for adjusting behavior as the phenotype under selection and check whether this affects our predictions.

In fact, it turns out that the evolutionary equilibrium for behavioral rules is not necessarily the same as that for fixed actions. Such a discrepancy has been highlighted in the context of parental provisioning of offspring, both for the conflict between two parents ([McNamara et al. 1999](#); [Taylor and Day 2004](#); [Johnstone and Hinde 2006](#)) and that between parent and offspring ([Kölliker 2003](#); [Smiseth et al. 2008](#); [Dobler and Kölliker 2009](#)). The standard approach to analyze these conflicts is to identify the best fixed action for one party given the fixed action of the other party, which examines conflict resolution solely on an evolutionary timescale and ignores the fact that both parties will normally be able to respond to each other's behavior in a flexible manner during the course of their interaction. More recent approaches have incorporated this behavioral flexibility by considering rules for responding instead of fixed actions, and they find that the stable level of provisioning differs from standard predictions ([McNamara et al. 1999](#); [Kölliker 2003](#); [Taylor and Day 2004](#); [Johnstone and Hinde 2006](#); [Smiseth et al. 2008](#); [Dobler and Kölliker 2009](#)). Thus, in order to predict adaptive behavior, it is necessary first to focus on the behavioral rule rather than the actions it produces.

Whether similar discrepancies exist in other contexts remains to be investigated, but we see no reason why this should be something specific to parental care ([McNamara et al. 1999](#); [Dobler and Kölliker 2009](#)). In cases where behavior is flexible and can be adjusted in response to the behavior of others, we should not just assume that the BSS and the ESS will coincide. This is an open question that we can answer with the appropriate type of analysis.

## Development of behavior

By focusing on selective optima and evolutionary trajectories toward those optima, behavioral ecologists tend to neglect how behavior develops during the lifetime of an individual (but see [Caro and Bateson 1986](#); [Bateson 1988, 2001](#); [Stamps 2003](#)). Development and function—the domains of two of [Tinbergen's \(1963\)](#) four “why” questions about

behavior—have traditionally been studied separately, but we believe that much could be gained by integrating these levels of analysis (see also [Laland et al. 2011](#)). By considering the evolved psychological mechanisms underlying behavior, we can begin to understand the behavioral changes that take place as individuals develop from an adaptive viewpoint and predict how behavior should change in the light of experience.

A dynamic model of aggression by [Fawcett and Johnstone \(2010\)](#) illustrates this approach. Fawcett and Johnstone evolved a rule for aggressive behavior dependent on past experience of victories and defeats, thereby allowing behavior to change within an individual's lifetime, as well as on an evolutionary timescale. They assumed that individuals are uncertain of their own strength, but that they gradually learn about this through their aggressive interactions with others. The model generated developmental trajectories of aggression under adaptive behavior at the evolutionary equilibrium, predicting that aggression should peak in young, naive individuals and decline thereafter as they gain experience. This pattern fits well with observed changes in aggressive behavior in humans from birth to adulthood ([Tremblay 2010](#)). The model's predictions of age-dependent behavior were a result of adaptive rules for learning from past experiences and could not have been generated using standard models of aggression.

There is great potential for more work in this area. Most optimality models (including those that allow for state dependence) take a simplified view of adaptive behavior, in which individuals immediately show appropriate responses to the situations they encounter and follow a predetermined strategy throughout their life, with changes occurring only through evolution. But these responses are in fact the product of underlying rules that play a critical role in development, causing behavior to change as an individual interacts with its environment. Viewed in this light, juvenile behavior is not simply an imperfectly formed version of adaptive behavior in adults; rather, both may result from the same adaptive psychological mechanisms. We encourage other researchers to study the ontogeny of behavior as a product of evolution ([Caro and Bateson 1986](#); [Bateson 2001](#)) by considering how the mechanisms underlying adaptive behavior orchestrate changes during development.

## Cognitive ecology

Behavioral phenotypes are a result of cognition—“the neuronal processes concerned with the acquisition, retention, and use of information” ([Dukas and Ratcliffe 2009](#)). The emerging field of cognitive ecology ([Dukas 1998](#); [Dukas and Ratcliffe 2009](#); [Shettleworth 2010](#)) deals with how cognition is shaped by evolution in a species' natural environment, but it is uncommon to incorporate the details of proximate mechanisms into optimality analyses of behavior (for notable exceptions, see [Kacelnik and Todd 1992](#); [Todd and Kacelnik 1993](#); [Brunner et al. 1996](#); [Stephens and Anderson 2001](#); [Stephens 2002](#); [Shapiro et al. 2008](#)).

Cognitive systems are full of biases. For example, some sensory receptors have biased sensitivity to particular stimuli; exaggerated representations of natural stimuli can elicit supernormal responses; discrimination between positive and negative stimuli is often affected by peak shift; some types of associations are learnt more readily than others; and responses to multiple stimuli may interact in complex ways ([Enquist and Arak 1998](#)). Rats ([Harding et al. 2004](#)), starlings ([Matheson et al. 2008](#)), and honeybees ([Bateson et al. 2011](#)) show apparently “optimistic” or “pessimistic” biases in how they respond to ambiguous stimuli. Likewise, at the level of decision making, animals are often impulsive, irrational, and

show response matching instead of maximizing their rewards (see above). From a functional viewpoint, these kinds of biases make little sense if we consider isolated behaviors; but if we consider the adaptive value of the psychological mechanisms responsible for those behaviors, then we might begin to understand why they have evolved (Haselton et al. 2009; McNamara and Houston 2009).

Hamblin and Giraldeau (2009) modeled foraging behavior in a producer–scrounger game (Barnard and Sibly 1981). Rather than following the standard approach of tracking frequency-dependent selection on producing and scrounging behavior (Giraldeau and Caraco 2000), they compared different learning mechanisms proposed to influence the use of these tactics and allowed the parameter values of each mechanism to evolve. The evolved values for the most successful mechanism showed an unexpected bias: the residual tendency for individuals to use the producer tactic when this went unrewarded was much higher than that for the scrounger tactic, which was close to zero (Hamblin and Giraldeau 2009). In practice, this implies that producing behavior should be difficult to extinguish, whereas the use of scrounging should be much more responsive to the level of reinforcement. The novel prediction of a learning bias is beyond the scope of standard fixed-action approaches to analyzing the producer–scrounger game; it is a direct consequence of modeling the learning mechanisms underlying behavior.

Similar insights have come from other studies which have considered the adaptive value of learning mechanisms. Groß et al. (2008) studied foraging behavior in a changing environment, and similar to Hamblin and Giraldeau (2009), they found that performance was critically affected by the tendency to keep sampling the pay-off from actions which are not currently favored. Over an extended time period, behavioral rules which stop sampling alternative actions perform poorly. Learning biases can also be predicted for contexts outside foraging. In Fawcett and Johnstone's (2010) model of aggression, learning about one's own strength drives winner and loser effects, in which aggressive tendencies are modulated by past experiences of victory and defeat. Winner and loser effects have been widely documented in the empirical literature (Hsu et al. 2006), but the model's novel prediction that these effects should change with age remains to be tested.

Developing a coherent body of theory to explain cognitive biases is an ambitious goal, and one that is a long way off yet. But we believe that our best hope of understanding such biases lies in studying the adaptive value of psychological mechanisms, not the individual behaviors those mechanisms produce.

## WHAT KINDS OF MECHANISMS SHOULD WE EXPECT?

As argued above, we should expect animals to have evolved psychological mechanisms which work well across a range of contexts (Gigerenzer et al. 1999; Hutchinson and Gigerenzer 2005; McNamara and Houston 2009) that share similar statistical properties (Todd and Gigerenzer 2007; Wilke and Barrett 2009). Behavior in any given situation may not be exactly optimal, but overall the animal should perform in an efficient way. This view is related to Simon's (1956) concepts of “satisficing” and “bounded rationality”, long popular among social scientists studying decision making in humans (Rubinstein 1998; Gigerenzer and Selten 2002; Kahneman 2003). However, in contrast to what we are advocating here, the bounded rationality perspective rarely adopts an evolutionary approach. When

assessing adaptive function, we need to ask what kinds of psychological mechanisms will be favored by natural selection (Stevens 2008; McNamara and Houston 2009). Some will be fixed rules for responding to stimuli (Hutchinson and Gigerenzer 2005), but where flexible behavior is concerned, they will often be partly shaped by learning. For the remainder of our article, we focus on rules for learning, because these have been the subject of most studies in this area and the findings have broad implications for behavior in a variety of contexts.

In the perspective we are advocating, adaptive behavior is determined not by the evolutionary stability of individual actions but the evolutionary stability of the underlying rules. It was Harley (1981) who first applied this concept to learning mechanisms when he searched for an evolutionarily stable (ES) learning rule (Table 1): a rule that, when adopted by all individuals in a population, cannot be outcompeted by any alternative rule used by a mutant individual. Harley claimed that the ES learning rule would be the one that leads most quickly to the ESS for fixed actions, and on this basis, he identified the “relative pay-off sum” (RPS) as a rule which came close to being evolutionarily stable (Harley 1981; Maynard Smith 1984). Under the RPS rule, an individual performs each action in its repertoire with a probability equal to the cumulative pay-off received for that action so far, relative to the total sum of pay-offs for all actions; it therefore approximates matching (see above) under many conditions (but not all; see Houston and Sumida 1987). Harley also included a “memory factor” which determines the weighting given to past events and a residual probability of continuing to perform an action in the absence of reward (see above). His conclusions were criticized by Houston (1983) and Houston and Sumida (1987), who showed that under some circumstances, fitness could be increased by deviating from matching. Tracy and Seaman (1995) corrected the mathematical details of Harley's analysis and showed that both the RPS rule and the ES learning rule tend to converge to matching, but they pointed out that many other rules might do the same (see also Selten and Hammerstein 1984). There are two important conclusions to draw from this debate. First, because matching does not uniquely specify behavior (Houston and McNamara 1981), it is not possible to infer the rule an animal is using simply from the observation of matching. Second, the fact that a given rule generates matching does not guarantee its evolutionary stability.

After waning somewhat, interest in the evolution of learning rules has recently been reignited by computer simulation studies (e.g. Beauchamp 2000; Groß et al. 2008; Hamblin and Giraldeau 2009; Buchkremer and Reinhold 2010) investigating the performance of the RPS rule and alternative learning rules, principally the perfect memory (PM) rule and the linear operator (LO) rule (for mathematical details of all these rules, see Hamblin and Giraldeau 2009). Interestingly, the RPS rule again seems to do best (Hamblin and Giraldeau 2009). Although we commend these attempts to study the adaptive value of learning rules, we believe that current approaches suffer from several limitations. Researchers have compared performance of a number of candidate rules, but these are selected with little theoretical justification, often simply using rules previously mentioned in the literature. The apparent success of the RPS rule does not preclude the existence of a vast number of alternative rules that would perform equally well, or even better (Selten and Hammerstein 1984). Cognitive psychologists have considered a greater diversity of rules in their work on heuristics for decision making in humans (Gigerenzer et al. 1999), but this research is only loosely based on

evolutionary considerations and typically uses arbitrary measures of performance. To give one example, [Todd and Miller \(1999\)](#) identified “adjust relative/2” as the best rule for learning during mate search, but this was based on the degree of assortative mating it produced, which is not necessarily something that would be optimized by selection. We feel this line of work would benefit from performance measures more closely tied to evolutionary fitness. What is critical is not how well a rule performs according to some subjectively chosen metric, but whether it will be favored by natural selection. Even in studies which have adopted an evolutionary approach, however, the parameter values for candidate rules are often chosen arbitrarily instead of being allowed to evolve (e.g. [Beauchamp 2000](#)), making it difficult to identify the best rule ([Lea and Dow 1984](#); [Hamblin and Giraldeau 2009](#)). Moreover, few studies actually pit rules directly against each other, instead comparing their performance across separate simulations. Because evolution proceeds through direct competition between variants, a truly evolutionary simulation should reflect this.

## FUTURE DIRECTIONS

We have criticized standard approaches to the study of behavior which consider the evolution of individual actions and highlighted the potential benefits of shifting our focus to the psychological mechanisms governing behavioral flexibility. It is important that we now define a clear way forward. To understand the evolution of learning rules, there are three key issues we need to address. First, we need to investigate the adaptive value of different possible rules from a theoretical point of view to predict which ones should evolve. Second, we need in-depth, empirical description of learning rules to identify how real animals learn in a range of different contexts. Third, we need to document individual variation in learning rules and measure the fitness consequences of this variation. Below we outline a research programme for tackling these issues.

### Which learning rules are favored by natural selection?

At a theoretical level, we recommend a more systematic approach to predicting which types of learning rules will be favored by natural selection. There are two key issues to consider when comparing the fitness of different rules. The first issue is that of evolutionary stability: does the widespread adoption of a given rule in a population guarantee that alternative rules cannot invade? Historically, most studies have focused on this question ([Harley 1984](#); [Houston 1983](#); [Maynard Smith 1984](#); [Selten and Hammerstein 1984](#); [Houston and Sumida 1987](#); [Tracy and Seaman 1995](#)). However, the stability criterion tells us nothing about whether evolution will lead the population to that equilibrium in the first place. A rule may be evolutionarily stable but not evolutionarily attractive, in which case a population that does not start by uniformly adopting that rule can never evolve to use it ([Nowak 1990](#)). So, the key questions we need to ask are: is a learning rule evolutionarily stable, and is it an evolutionary attractor?

It is also important to evaluate these rules in a wider variety of circumstances than is currently done. It is natural that early work on learning rules placed emphasis on simple scenarios such as the ideal free distribution and the two-armed bandit problem ([Maynard Smith 1984](#)). However, as [Houston et al. \(2007\)](#) have pointed out, to derive more realistic predictions, we will have to confront our models with additional complications such as environmental fluctuation

and information about future rewards. To this, we would add that it is crucial that more models of learning rules take spatial processes into account ([Valcu and Kempenaers 2010](#)). Illustrating this, work on the spatial structure of foragers playing a producer–scrounger game has shown that a simple, spatially explicit rule for behavior can lead to complex behavioral patterns of the kind seen in laboratory studies of this game ([Hamblin and Giraldeau unpublished](#)).

With a view to testing hypotheses about the adaptiveness of learning rules, the aim should be to create models of animal behavior which drive empirical research. Although it is generally preferable to have mathematical models which can be solved analytically, adding crucial biological details can sometimes make analytical solutions intractable. We believe that other theoretical techniques have much to offer and therefore welcome the recent trend toward a heavier reliance on numerical analysis and simulation. Alongside more well-established tools such as stochastic dynamic programming ([Houston and McNamara 1999](#)), the use of genetic algorithms ([Sumida et al. 1990](#); [Hamblin and Hurd 2007](#); [Ruxton and Beauchamp 2008](#)) and individual-based simulations ([DeAngelis and Mooij 2005](#)) has great potential for modeling animal behavior through time and space. Genetic algorithms allow us to simulate the emergence and evolutionary dynamics of behavioral rules, revealing both their evolutionary stability and evolutionary attractiveness. By pitting a large number of rules against each other, we can explore the resulting population dynamics, and by coupling genetic algorithms with individual-based simulations, we can implement more realistic depictions of learning behavior which are spatially explicit. Using this approach, future research can identify the key properties of successful learning rules across multiple contexts, as a guide to empirical investigation. The behavior those rules would produce in specific situations can then be compared directly with observed behavior in the field and laboratory.

### How do real animals learn?

Thanks to many decades of research in experimental psychology, we have a fairly good understanding of how animals learn in a foraging context. As we have seen, this has highlighted some intriguing cases in which animals appear not to behave adaptively, perhaps because they are confronted with artificial situations in the laboratory for which their evolutionary history has not prepared them ([McNamara and Houston 1980](#); [Houston and McNamara 1989](#); [Houston et al. 2007](#); [Houston 2009](#); [Pavlic and Passino 2010](#)). Nonetheless, there exists a wealth of data on how animals sample available options, how they respond to positive reinforcement and punishment and how they remember past experiences. For contexts outside foraging, where rewards are not measured in terms of food, we know far less. There has been some work on learning in an aggressive context, mainly in terms of winner and loser effects ([Hsu et al. 2006](#)). In most study systems, loser effects seem to be more powerful than winner effects, but we lack a good explanation for why this bias exists. For other interactive contexts, such as mating and cooperating, we know next to nothing about the ways in which animals learn. This is an unfortunate gap in our understanding, given that learning is likely to play a key role in the development of behavior in these domains too. Interestingly, one of the few clear experimental tests of the iterated prisoner’s dilemma found that blue jays (*Cyanocitta cristata*) had difficulty learning to cooperate in this contrived laboratory situation, perhaps because of an impulsive tendency to choose short-term over larger long-term rewards ([Clements and Stephens 1995](#); [Stevens and Stephens 2004](#); but see [St-Pierre et al. 2009](#)).



Likewise, learning rules like the RPS rule have been studied mainly in foraging situations, for example patch choice (Harley 1981), the multi-armed bandit (Harley 1981; Groß et al. 2008; Buchkremer and Reinhold 2010) and the producer–scrounger game (Beauchamp 2000; Hamblin and Giraldeau 2009). As illustrated by Harley (1981) these simple learning rules can easily be applied in other domains, suggesting that there may well be some very general principles of learning that are adaptive across a range of contexts. Just as a food reward acts as a positive reinforcer in a foraging context, it is straightforward to think of stimuli that might act as positive reinforcers in other contexts: defeating an opponent in an aggressive context, copulation in a mating context, shared success in a cooperative context. It could be that similar learning rules govern the way animals respond to all of these reinforcers, but currently we lack the data to answer this question. A research priority should be to study in detail how animals learn in non-foraging situations.

### How do individuals differ in learning rules?

So far, our discussions concerning the mechanisms underlying behavioral flexibility have assumed that these are common to all individuals of a given species, and perhaps even evolutionarily conserved across many species. This reflects standard approaches to studying animal behavior. In experimental psychology, for example, research focuses on general principles of learning—processes such as habituation, sensitization, and conditioning—and emphasizes species-typical patterns rather than individual differences (Barker and Katz 2003). The widespread assumption is that the rules of associative learning in one animal do not differ from those in another. Similarly, when it comes to theoretical studies on learning rules, the quest is often to find the optimal learning rule or the one that takes a population most quickly to an ESS, from which it cannot be displaced by any alternative rule (Harley 1981; Maynard Smith 1984). This assumes that one type of learning rule is the most adaptive and will exclude all others. However, there is good reason to believe this may not always be so.

In an evolutionary model of individual differences in responsiveness, Wolf et al. (2008) found that responsiveness fails to spread to the whole population in a patch-choice situation. They considered a scenario in which members of a population disperse to two foraging patches that differ in quality. Unresponsive individuals choose patches randomly, whereas responsive individuals sample first and then choose to go to the patch which yields the highest pay-off. As responsive individuals spread, their success declines as a result of overcrowding at the initially better patch, until at some point they do no better than individuals that choose randomly. Assuming some cost of responsiveness (e.g. the time or energy invested in sampling), the result is an ES combination of responsive and unresponsive individuals.

More recently, variable versus fixed behavior was explored in the context of the producer–scrounger game (Dubois et al. 2010). The standard version of this game considers only fixed phenotypes—that is, individuals are either pure producers or pure scroungers—and the predicted ESS in most cases is a stable mixture of producers and scroungers (Barnard and Sibly 1981; Giraldeau and Caraco 2000; Giraldeau and Dubois 2008). When foraging conditions change, favoring a different ratio of the two phenotypes, frequency-dependent selection moves the population toward the new ESS over several generations. Dubois et al. (2010) introduced a mutant which can directly sample the pay-off from producing and scrounging and flexibly adjust its use of these behaviors so as to maximize its pay-offs. This flexible mutant initially does

better than the fixed producers and scroungers, so it begins to spread; but in doing so it changes the pay-offs obtained by the fixed phenotypes. Once it is sufficiently numerous in the population, its flexibility buffers the fixed phenotypes against selection caused by changes in foraging conditions. As a result, the flexible phenotype never spreads to fixation, and in fact it always remains in the minority regardless of the size of the population or the extent to which foraging conditions change (Dubois et al. 2010).

Both of these studies illustrate that the learning or flexibility of some individuals will critically affect the pay-offs of flexibility in others. They make the intriguing prediction that populations of animals will consist of a mixture of types which differ in their responsiveness or learning ability: some individuals are malleable, whereas others are rigid in their behavior; some individuals are fast learners, whereas others are slow. One could therefore imagine different learning rules coexisting within the same population as a stable polymorphism. Alternatively, perhaps conspecifics share the same basic learning rule but are polymorphic in its precise specifications (i.e. in the parameter values), for example by weighting past events differently or differing in their sensitivity to reinforcement. Thus, differences in learning ability need not represent random, nonadaptive variation giving rise to “bright” and “dim” subjects; they might be actively maintained by frequency-dependent selection.

We propose that such differences in learning ability may underlie the variation in “personalities” documented in a growing number of species (Réale et al. 2007). Learning is associated with a range of behavior patterns which show consistent between-individual variation, most notably exploration. In great tits (*Parus major*), for example, bold individuals explore quickly but superficially and are slow to respond to changes in their environment, whereas shy individuals explore more thoroughly and are much more responsive to environmental change (Verbeek et al. 1994). Such differences in responsiveness appear to be a key component of behavioral variation (Dingemans et al. 2010). But whereas the literature on animal personalities addresses individual differences in behavior, the approach we advocate emphasizes individual differences in the underlying learning rules. That is, individuals may differ in a variety of personality traits as a direct result of differences in how they learn. The possibility of polymorphic learning within populations may provide a much-needed framework for the study of individual differences in behavior. A key challenge for the future is to document the heritability and the fitness consequences of differences in learning ability.

### CONCLUDING REMARKS

Others before us (Houston 1987; Dukas 1998; Hutchinson and Gigerenzer 2005; McNamara and Houston 2009; Dukas and Ratcliffe 2009; Shettleworth 2010; Laland et al. 2011) have called for a greater integration of psychological and biological approaches to studying animal behavior. These calls have largely gone unheeded; lip service has been paid to the benefits of a more integrated approach, but rarely has this been embraced in practice. Rather than studying function and mechanism in parallel, we should be considering psychological mechanisms from an adaptive viewpoint, in an attempt to understand how they have evolved.

We are not suggesting that the behavioral gambit should be abandoned forthwith; rather, we are questioning its blind acceptance. Our concern is that most behavioral ecologists seem largely unaware that they are making an assumption that behavior is unconstrained by the psychological mechanisms governing flexibility. There appears to be a

widespread belief that the evolution of decision rules will generate the same solutions as the evolution of fixed actions, so that we need not worry about the details of those rules. Does the behavioral gambit stand up to empirical scrutiny? In many cases it may hold, although behavioral data show that sometimes it fails; animals do not behave optimally in all situations they encounter. But in any case, the point we wish to make is that the validity of the behavioral gambit is an empirical question, and it is important that we address it directly. We urge behavioral ecologists to turn their attention to the evolution of decision mechanisms, as multipurpose rules which are capable of providing effective solutions to a wide range of problems.

## FUNDING

This work was supported by the European Commission (FP7-PEOPLE-2009-IOF, grant number 252618 to T.W.F.); the Natural Sciences & Engineering Research Council of Canada (CGS-D grant to S.H.); and the European Research Council (Advanced Grant number 250209 to Alasdair Houston).

We thank Neeltje Boogert, Rob Brooks, Jonathan Grose, Alasdair Houston, Pete Hurd, Jeff Stevens, and an anonymous referee for valuable feedback on earlier versions of the manuscript.

## REFERENCES

- Aw JM, Holbrook RI, Burt de Perera T, Kacelnik A. 2009. State-dependent valuation learning in fish: banded tetras prefer stimuli associated with greater past deprivation. *Behav Processes*. 81:333–336.
- Barker L, Katz JS. 2003. Animal learning and animal cognition. In: Davis SF, editor. *Handbook of research methods in experimental psychology*. Oxford: Blackwell Publishing. p. 241–262.
- Barnard CJ, Sibly RM. 1981. Producers and scroungers: a general model and its application to captive flocks of house sparrows. *Anim Behav*. 29:543–550.
- Bateson M. 2002. Context-dependent foraging choice in risk-sensitive starlings. *Anim Behav*. 64:251–260.
- Bateson M, Desire S, Gartside SE, Wright GA. 2011. Agitated honeybees exhibit pessimistic cognitive biases. *Curr Biol*. 21:1070–1073.
- Bateson M, Healy SD. 2005. Comparative evaluation and its implications for mate choice. *Trends Ecol Evol*. 20:659–664.
- Bateson M, Healy SD, Hurlly TA. 2003. Context-dependent foraging decisions in rufous hummingbirds. *Proc R Soc B*. 270:1271–1276.
- Bateson P. 1988. The active role of behaviour in evolution. In: Ho MW, Fox S, editors. *Evolutionary processes and metaphors*. Chichester, UK: John Wiley. p. 191–207.
- Bateson P. 2001. Behavioral development and Darwinian evolution. In: Oyoma S, Griffiths PE, Gray RD, editors. *Cycles of contingency: developmental systems and evolution*. Cambridge (MA): MIT Press. p. 149–166.
- Beauchamp G. 2000. Learning rules for social foragers: implications for the producer–scrounger game and ideal free distribution theory. *J Theor Biol*. 207:21–35.
- Brunner D, Kacelnik A, Gibbon J. 1996. Memory for inter-reinforcement interval variability and patch departure decisions in the starling, *Sturnus vulgaris*. *Anim Behav*. 51:1025–1045.
- Buchkremer EM, Reinhold K. 2010. The emergence of variance-sensitivity with successful decision rules. *Behav Ecol*. 21:576–583.
- Bull JJ, Wang I-N. 2010. Optimality models in the age of experimental evolution and genomics. *J Evol Biol*. 23:1820–1838.
- Capaldi ED, Myers DE, Campbell DH, Sheffer JD. 1983. Conditioned flavor preferences based on hunger level during original flavor exposure. *Anim Learn Behav*. 11:107–115.
- Caro TM, Bateson P. 1986. Organization and ontogeny of alternative tactics. *Anim Behav*. 34:1483–1499.
- Cheverud JM. 1988. A comparison of genetic and phenotypic correlations. *Evolution*. 42:958–968.
- Clements KC, Stephens DW. 1995. Testing models of non-kin cooperation: mutualism and the Prisoner's Dilemma. *Anim Behav*. 50:527–535.
- Davison M, McCarthy D. 1988. *The matching law: a research review*. Hillsdale (NJ): Lawrence Erlbaum Associates.
- Dawkins R. 1976. *The selfish gene*. Oxford: Oxford University Press.
- DeAngelis DL, Mooij WM. 2005. Individual-based modeling of ecological and evolutionary processes. *Annu Rev Ecol Evol Syst*. 36:147–168.
- Dingemanse NJ, Kazem AJN, Réale D, Wright J. 2010. Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol Evol*. 25:81–89.
- Dobler R, Kölliker M. 2009. Behavioural attainability of evolutionarily stable strategies in repeated interactions. *Anim Behav*. 77:1427–1434.
- Dubois F, Morand-Ferron J, Giraldeau L-A. 2010. Learning in a game context: strategy choice by some keeps learning from evolving in others. *Proc R Soc B*. 277:3609–3616.
- Dukas R (editor). 1998. *Cognitive ecology: the evolutionary ecology of information processing and decision making*. Chicago (IL): University of Chicago Press.
- Dukas R, Ratcliffe JM (editors). 2009. *Cognitive ecology II*. Chicago (IL): University of Chicago Press.
- Enquist M, Arak A. 1998. Neural representation and the evolution of signal form. In: Dukas R, editor. *Cognitive ecology: the evolutionary ecology of information processing and decision making*. Chicago (IL): University of Chicago Press. p. 21–88.
- Fawcett TW, Johnstone RA. 2010. Learning your own strength: winner and loser effects should change with age and experience. *Proc R Soc B*. 277:1427–1434.
- Gibbon J. 1977. Scalar expectancy theory and Weber's law in animal timing. *Psychol Rev*. 84:279–325.
- Gigerenzer G, Selten R (editors). 2002. *Bounded rationality: the adaptive toolbox*. Cambridge (MA): MIT Press.
- Gigerenzer G, Todd PM, ABC Research Group (editors). 1999. *Simple heuristics that make us smart*. New York: Oxford University Press.
- Giraldeau L-A, Caraco T. 2000. *Social foraging theory*. Princeton (NJ): Princeton University Press.
- Giraldeau L-A, Dubois F. 2008. Social foraging and the study of exploitative behavior. *Adv Stud Behav*. 38:59–104.
- Grafen A. 1984. Natural selection, kin selection and group selection. In: Krebs JR, Davies NB, editors. *Behavioural ecology: an evolutionary approach*, 2nd edition. Oxford, UK: Blackwell Scientific Press. p. 62–84.
- Groß R, Houston AI, Collins EJ, McNamara JM, Dechaume-Montcharmont F-X, Franks NR. 2008. Simple learning rules to cope with changing environments. *J R Soc Interface*. 5:1193–1202.
- Hadfield JD, Nutall A, Osorio D, Owens IPF. 2007. Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *J Evol Biol*. 20:549–557.
- Hamblin S, Giraldeau L-A. 2009. Finding the evolutionarily stable learning rule for frequency-dependent foraging. *Anim Behav*. 78:1343–1350.
- Hamblin S, Hurd PL. 2007. Genetic algorithms and non-ESS solutions to game theory models. *Anim Behav*. 74:1005–1018.
- Hammerstein P. 1998. What is evolutionary game theory? In: Dugatkin LA, Reeve HK, editors. *Game theory and animal behavior*. New York: Oxford University Press. p. 3–15.
- Harding EJ, Paul ES, Mendl M. 2004. Cognitive bias and affective state. *Nature*. 427:312.
- Harley CB. 1981. Learning the evolutionarily stable strategy. *J Theor Biol*. 89:611–633.
- Haselton MG, Bryant GA, Wilke A, Frederick DA, Galperin A, Frankenhuis WE, Moore T. 2009. Adaptive rationality: an evolutionary perspective on cognitive bias. *Soc Cogn*. 27:733–763.
- Henly SE, Ostdiek A, Blackwell E, Knutic S, Dunlap AS, Stephens DW. 2008. The discounting-by-interruptions hypothesis: model and experiment. *Behav Ecol*. 19:154–162.
- Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R. 2001. In search of *Homo economicus*: behavioral experiments in 15 small-scale societies. *Am Econ Rev*. 91:73–78.
- Herrnstein RJ. 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav*. 4:267–272.
- Herrnstein RJ. 1990. Rational choice theory: necessary but not sufficient. *Am Psychol*. 45:356–367.



- Herrnstein RJ, Heyman GM. 1979. Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *J Exp Anal Behav*. 31:209–233.
- Herrnstein RJ, Vaughan WJ. 1980. Melioration and behavioral allocation. In: Staddon JER, editor. *Limits to action*. New York: Academic Press. p. 143–176.
- Heyman GM, Herrnstein RJ. 1986. More on concurrent interval-ratio schedules: a replication and review. *J Exp Anal Behav*. 46:331–351.
- Hori M. 1993. Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science*. 260:216–219.
- Houston AI. 1983. Comments on “learning the evolutionarily stable strategy”. *J Theor Biol*. 105:175–178.
- Houston AI. 1987. The control of foraging decisions. In: Commons ML, Kacelnik A, Shettleworth SJ, editors. *Quantitative analyses of behavior: foraging*, vol. 6. Hillsdale (NJ): Lawrence Erlbaum Associates. p. 41–61.
- Houston AI. 2009. Flying in the face of nature. *Behav Processes*. 80:295–305.
- Houston AI, McNamara JM. 1981. How to maximize reward rate in two variable-interval paradigms. *J Exp Anal Behav*. 35:367–396.
- Houston AI, McNamara JM. 1989. The value of food: effects of open and closed economies. *Anim Behav*. 37:546–562.
- Houston AI, McNamara JM. 1999. *Models of adaptive behaviour: an approach based on state*. Cambridge, UK: Cambridge University Press.
- Houston AI, McNamara JM, Steer MD. 2007. Do we expect natural selection to produce rational behaviour? *Phil Trans R Soc B*. 362:1531–1543.
- Houston AI, Sumida BH. 1987. Learning rules, matching and frequency dependence. *J Theor Biol*. 126:289–308.
- Hsu Y, Earley RL, Wolf LL. 2006. Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biol Rev*. 81:33–74.
- Hutchinson JMC, Gigerenzer G. 2005. Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behav Processes*. 69:97–124.
- Johnstone RA, Hinde CA. 2006. Negotiation over offspring care—how should parents respond to each other’s efforts? *Behav Ecol*. 17:818–827.
- Kacelnik A. 2006. Meanings of rationality. In: Hurley S, Nudds M, editors. *Rational animals?* Oxford, UK: Oxford University Press. p. 87–106.
- Kacelnik A, Todd IA. 1992. Psychological mechanisms and the Marginal Value Theorem: effect of variability in travel time on patch exploitation. *Anim Behav*. 43:313–322.
- Kagel JH, Green L, Caraco T. 1986. When foragers discount the future: constraint or adaptation? *Anim Behav*. 34:271–283.
- Kahneman D. 2003. Maps of bounded rationality: psychology for behavioral economics. *Am Econ Rev*. 93:1449–1475.
- Kölliker M. 2003. Estimating mechanisms and equilibria for offspring begging and parental provisioning. *Proc R Soc B (Suppl)*. 270:S110–S113.
- Krebs JR, Davies NB. 1997. The evolution of behavioural ecology. In: Krebs JR, Davies NB, editors. *Behavioural ecology: an evolutionary approach*, 3rd edition. Oxford, UK: Blackwell Science. p. 3–12.
- Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T. 2011. Cause and effect in biology revisited: is Mayr’s proximate-ultimate distinction still useful? *Science*. 334:1512–1516.
- Lea SEG, Dow SM. 1984. Optimization and flexibility. *Behav Brain Sci*. 7:110–111.
- Loewenstein G. 2000. Emotions in economic theory and economic behavior. *Am Econ Rev*. 90:426–432.
- Marsh B, Schuck-Paim C, Kacelnik A. 2004. Energetic state during learning affects foraging choices in starlings. *Behav Ecol*. 15:396–399.
- Matheson SM, Asher L, Bateson M. 2008. Larger, enriched cages are associated with ‘optimistic’ response biases in captive European starlings (*Sturnus vulgaris*). *Appl Anim Behav Sci*. 109:374–383.
- Maynard Smith J. 1982. *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
- Maynard Smith J. 1984. Game theory and the evolution of behaviour. *Behav Brain Sci*. 7:95–101.
- Mazur JE. 2000. Tradeoffs among delay, rate, and amount of reinforcement. *Behav Processes*. 49:1–10.
- McNamara JM, Gasson CE, Houston AI. 1999. Incorporating rules for responding into evolutionary games. *Nature*. 401:368–371.
- McNamara JM, Houston AI. 1980. The application of statistical decision theory to animal behaviour. *J Theor Biol*. 85:673–690.
- McNamara JM, Houston AI. 2009. Integrating function and mechanism. *Trends Ecol Evol*. 24:670–675.
- Morand-Ferron J, Giraldeau L-A. 2010. Learning behaviorally stable solutions to producer–scrounger games. *Behav Ecol*. 21:343–348.
- Mottley K, Giraldeau L-A. 2000. Experimental evidence that group foragers can converge on predicted producer–scrounger equilibria. *Anim Behav*. 60:341–350.
- Nowak M. 1990. An evolutionarily stable strategy may be inaccessible. *J Theor Biol*. 142:237–241.
- Parker GA, Maynard Smith J. 1990. Optimality theory in evolutionary biology. *Nature*. 348:27–33.
- Pavlic TP, Passino KM. 2010. When rate maximization is impulsive. *Behav Ecol Sociobiol*. 64:1255–1265.
- Persky J. 1995. Retrospectives: the ethology of *Homo economicus*. *J Econ Perspect*. 9:221–231.
- Pompilio L, Kacelnik A, Behmer ST. 2006. State-dependent learned valuation drives choice in an invertebrate. *Science*. 311:1613–1615.
- Puts DA. 2010. The psychologic gambit declined—a review of “Endocrinology of Social Relationships”. *Evol Hum Behav*. 31:304–308.
- Rachlin H, Green L. 1972. Commitment, choice and self-control. *J Exp Anal Behav*. 17:15–22.
- Réale D, Festa-Bianchet M. 2000. Quantitative genetics of life-history traits in a long-lived wild mammal. *Heredity*. 85:593–603.
- Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007. Integrating animal temperament within ecology and evolution. *Biol Rev*. 82:291–318.
- Roff DA. 1995. The estimation of genetic correlations from phenotypic correlations: a test of Cheverud’s conjecture. *Heredity*. 74:481–490.
- Rubinstein A. 1998. *Modeling bounded rationality*. Cambridge (MA): MIT Press.
- Ruxton GD, Beauchamp G. 2008. The application of genetic algorithms in behavioural ecology, illustrated with a model of anti-predator vigilance. *J Theor Biol*. 250:435–448.
- Schuck-Paim C, Pompilio L, Kacelnik A. 2004. State-dependent decisions cause apparent violations of rationality in animal choice. *PLoS Biol*. 2:e402.
- Selten R, Hammerstein P. 1984. Gaps in Harley’s argument on evolutionarily stable learning rules and in the logic of “tit for tat”. *Behav Brain Sci*. 7:115–116.
- Shafir S, Waite TA, Smith BH. 2002. Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behav Ecol Sociobiol*. 51:180–187.
- Shapiro MS, Siller S, Kacelnik A. 2008. Simultaneous and sequential choice as a function of reward delay and magnitude: normative, descriptive and process-based models tested in the European starling (*Sturnus vulgaris*). *J Exp Psychol Anim Behav Process*. 34:75–93.
- Shettleworth SJ. 2010. *Cognition, evolution, and behavior*, 2nd edition. New York: Oxford University Press.
- Shiv B, Loewenstein G, Bechara A, Damasio H, Damasio AR. 2005. Investment behavior and the negative side of emotion. *Psychol Sci*. 16:435–439.
- Shuster SM, Wade MJ. 1991. Equal mating success among male reproductive strategies in a marine isopod. *Nature*. 350:608–610.
- Simon HA. 1956. Rational choice and the structure of the environment. *Psych Rev*. 63:129–138.
- Sinervo B, Lively CM. 1996. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature*. 380:240–243.
- Smiseth PT, Wright J, Kölliker M. 2008. Parent–offspring conflict and co-adaptation: behavioural ecology meets quantitative genetics. *Proc R Soc B*. 275:1823–1830.
- Staddon JER, Cerutti DT. 2003. Operant conditioning. *Annu Rev Psychol*. 54:115–144.
- Stamps J. 2003. Behavioural processes affecting development: Tinbergen’s fourth question comes of age. *Anim Behav*. 66:1–13.
- Stephens DW. 2002. Discrimination, discounting and impulsivity: a role for an informational constraint. *Phil Trans R Soc B*. 357:1527–1537.
- Stephens DW, Anderson D. 2001. The emergence of variance-sensitivity with successful decision rules. *Behav Ecol*. 21:576–583.

- Stephens DW, Clements KC. 1998. Game theory and learning. In: Dugatkin LA, Reeve HK, editors. *Game theory and animal behavior*. New York: Oxford University Press. p. 239–260.
- Stephens DW, Kerr B, Fernández-Juricic E. 2004. Impulsiveness without discounting: the ecological rationality hypothesis. *Proc R Soc B*. 271:2459–2465.
- Stevens JR. 2008. The evolutionary biology of decision making. In: Engel C, Singer W, editors. *Better than conscious? Decision making, the human mind, and implications for institutions*. Cambridge (MA): MIT Press. p. 285–304.
- Stevens JR, Stephens DW. 2004. The economic basis of cooperation: tradeoffs between selfishness and generosity. *Behav Ecol*. 15:255–261.
- St-Pierre A, Larose K, Dubois F. 2009. Long-term social bonds promote cooperation in the iterated Prisoner's Dilemma. *Proc R Soc B*. 276:4223–4228.
- Sumida BH, Houston AI, McNamara JM, Hamilton WD. 1990. Genetic algorithms and evolution. *J Theor Biol*. 147:59–84.
- Taylor PD, Day T. 2004. Stability in negotiation games and the emergence of cooperation. *Proc R Soc B*. 271:669–674.
- Tinbergen N. 1963. On aims and methods of ethology. *Z Tierpsychol*. 20:410–433.
- Todd IA, Kacelnik A. 1993. Psychological mechanisms and the Marginal Value Theorem: dynamics of scalar memory for travel time. *Anim Behav*. 46:765–775.
- Todd PM, Gigerenzer G. 2007. Environments that make us smart: ecological rationality. *Curr Dir Psychol Sci*. 16:167–171.
- Todd PM, Miller GF. 1999. From pride and prejudice to persuasion: satisficing in mate search. In: Gigerenzer G, Todd PM, ABC Research Group, editors. *Simple heuristics that make us smart*. New York: Oxford University Press. p. 287–308.
- Tracy ND, Seaman JW Jr. 1995. Properties of evolutionarily stable learning rules. *J Theor Biol*. 177:193–198.
- Tremblay RE. 2010. Developmental origins of disruptive behaviour problems: the 'original sin' hypothesis, epigenetics and their consequences for prevention. *J Child Psychol Psychiatry*. 51:341–367.
- Valcu M, Kempenaers B. 2010. Spatial autocorrelation: an overlooked concept in behavioral ecology. *Behav Ecol*. 21:902–905.
- Verbeek MEM, Drent PJ, Wiepkema PR. 1994. Consistent individual differences in early exploratory behaviour of male great tits. *Anim Behav*. 48:1113–1121.
- Weissing FJ. 1996. Genetic versus phenotypic models of selection: can genetics be neglected in a long-term perspective? *J Math Biol*. 34:533–555.
- Wilke A, Barrett HC. 2009. The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evol Hum Behav*. 30:161–169.
- Wolf M, van Doorn GS, Weissing FJ. 2008. Evolutionary emergence of responsive and unresponsive personalities. *Proc Natl Acad Sci USA*. 105:15825–15830.