

# Six Reasons for Invoking Evolution in Decision Theory

Peter Hammerstein and Jeffrey R. Stevens

It is in biology and psychology that economists and other social scientists will find the premises needed to fashion more predictive models.—E. O. Wilson (1998:206)

In 2012, a “Mega Millions” lottery in the United States set a world record with a jackpot of \$656 million. The estimated odds of winning the jackpot were 1 in over 175,000,000. Despite the terrible odds, this lottery generated a frenzy of ticket purchases: 650 million tickets were sold in three days. Why would so many people gamble on such poor odds?

On a cold winter morning, a young black-capped chickadee (*Poecile atricapillus*) awakens from sleep, very hungry. She faces the choice of flying off to an area with a low but consistent supply of seeds or an area with the possibility of many or no seeds (Barkan 1990). Which should she choose?

In both of these examples, a decision must be made on whether to gamble on a risky option or stick with a safe option. How do organisms make these decisions? The study of games and decisions has long been guided by a philosophical discourse on concepts of rationality and their implications. This discourse has led to a large body of mathematical work and kept generations of researchers busy, but few serious attempts have been made to understand decision making in the real world. Over the last decades, however, decision theory has moved toward the sciences and developed its “taste for the facts.” Research is now guided by experimental economics, cognitive psychology, behavioral biology, and—most recently—neuroscience. Despite the increasingly empirical leanings of decision science, the explanatory power of evolutionary theory has been neglected. We proposed this Strüngmann Forum to rectify this oversight, with the goal of initiating an alternative to the existing axiom-based decision theory by developing a theory of decision making founded on evolutionary principles.

Tackling this task requires a broad disciplinary base, and thus we assembled a group of researchers from a diverse range of fields—including evolutionary biology; cognitive, evolutionary, comparative, and developmental psychology;

neuroscience; computer science; economics, and philosophy—to approach this goal.

### **Toward a Darwinian Decision Theory**

In 1654, Blaise Pascal and Pierre Fermat laid the foundation for our current theory of decision making by developing the notion of *expected value* (Daston 1995). Accordingly, a reasonable organism need only estimate the value of an outcome, weight this value by the probability of the outcome occurring, and choose the outcome offering the highest expectation. This view was formalized mathematically in the 20th century, but for the most part it has remained unchanged. Despite the beauty of the theory, its descriptive validity has been plagued by contradictory data. In numerous cases, humans and other animals systematically violate the predictions made by decision theory (Kahneman and Tversky 2000; Gigerenzer et al. 1999; Rosati and Stevens 2009): They seek risks when the theory predicts that they should avoid them. They prefer immediate rewards when they should wait. They place a premium on items in their possession when trading offers a better deal. In short, decision theory all too often fails to predict behavior.

In the rush to catalog violations of decision theory, economists and psychologists failed to propose a viable alternative to the existing theory. When they did propose alternatives, they simply patched up the existing theory (e.g., prospect theory; see Kahneman and Tversky 1979). Whenever a new violation appeared, a new patch was placed over the old theory. Repairing the old theory, however, poses two problems. First, having too many patches greatly complicates the theory and suggests that the foundation of the theory is flawed. Second, the original decision theory and its successors ignore two factors critical to understanding decision making: evolution and cognitive mechanisms.

The current mismatch between decision theory and data highlights the benefits of starting afresh in our study of decision making. For the purposes of this Forum, we anticipated initiating a *theory of decision making that rests on an evolutionary foundation* by building from first principles. Note that this theory was intended to encompass not only human decision making but rather a theory of decision making that transcends a particular species and explores the general principles of how evolution can generate decision-making agents, biological or artificial.

This Strüngmann Forum assembled experts from varied disciplines to integrate a careful understanding of evolution with precise models of cognitive mechanisms. Our aim was to develop a springboard for the construction of a *Darwinian decision theory* from which deeper insights could be gained about the functionality of cognitive design. To this end, we discussed four key components of a Darwinian approach to decision making: (a) understanding the origins of decision mechanisms, (b) exploring why these mechanisms are robust,

(c) accounting for variation between and within individuals, and (d) investigating the pressures of social life on decision making. In this introductory chapter, we explore a few general themes that draw on discussions across these four areas. In particular, we enumerate six of the many reasons why a Darwinian approach could aid our understanding of decision making.

### Reasons for Invoking Evolution in Decision Theory

What would a Darwinian decision maker look like? Would she make transitive, consistent, logical decisions? Perhaps. Chater (this volume) refers to this as formal rationality (i.e., an emphasis on consistency in decision making). It seems unlikely, however, that natural selection would favor transitivity, consistency, and logic per se. Selection should favor *successful* rather than *consistent* decisions. While consistency may pay off in some situations, it may in others provide no benefits or even be costly when behavioral variation is favored (Brown et al., this volume). This idea maps onto Chater's notion of substantive rationality, which refers to decisions that require "an external standard, against which the quality of a decision can be measured." A Darwinian approach would provide substantive rationality because evolutionary fitness provides the ultimate external standard that selects for successful decision mechanisms (Stevens 2008; McNamara and Houston 2009; Hagen et al., this volume). Bearing in mind the evolutionary version of substantive rationality, we can begin to develop general principles of Darwinian decision making.

### Adaptive Specialization

From a biological point of view, the evolved decision mechanisms of a given species are expected to operate very efficiently in those environments that typically occurred during the evolutionary past of the species. Since a high degree of efficiency cannot be achieved by a general-purpose device, these mechanisms have likely been tuned to the environmental circumstances under which they have been used. This *adaptive specialization* is a key reason for "invoking Darwin" in the study of decision making.<sup>1</sup>

Though not without controversy (see Bolhuis and Macphail 2001), the notion of adaptive specialization in animal cognition and decision making is widespread (Shettleworth 2000). From spatial memory in caching birds (Kamil et al. 1994) to temporal preferences in primates (Rosati et al. 2007), we observe examples of cognitive abilities and decisions that differ among

---

<sup>1</sup> A Darwinian account is not restricted to adaptive explanations of behavior. A comprehensive Darwinian account includes influences of genetic drift, mutation, gene flow, and phylogenetic history. Though we do not consider these alternative mechanisms and constraints on evolution here, phylogenetic history is critical to the comparative approach taken by many decision scientists.

phylogenetically closely related species in ways that match the adaptive problems faced by those species in their natural environments. We, therefore, have evidence of adaptive specialization in animals.

Evolutionary psychologists use the concept of modularity to highlight the adaptive specialization of cognition in humans. Kurzban (this volume) argues that, rather than using Fodor's (1983) very strict and multifaceted definition of a cognitive module, the *evolutionary module* is defined by functional specialization. For instance, one of the classic examples proposed for evolutionary modules is cheater detection. Cosmides et al. (2010) argue that although people have difficulty solving abstract reasoning problems, they succeed when the problem is couched as an evolutionarily relevant situation, namely detecting cheaters. The authors contend that our minds did not evolve to solve content-free, abstract logical puzzles using *P*'s and *Q*'s. We did evolve, however, to avoid playing the sucker. When the logical puzzle is phrased as a social exchange, our "cheater-detection module" kicks in to solve the problem. Kurzban acknowledges the controversy surrounding the notion of an evolved module and provides justification for its utility.

In the evolutionary psychology tradition, Cosmides and Tooby (1994) liken the modular brain to a Swiss army knife. Analogously, in the judgment and decision-making field, Gigerenzer et al. (1999) endorse the "adaptive toolbox" approach of investigating the cognitive mechanisms used in specific decision contexts. Based on insights from Simon's (1956) concept of bounded rationality, the adaptive toolbox approach emphasizes fast and frugal heuristics that organisms use to solve adaptive problems. In the biological tradition, these heuristics are called "rules of thumb" (Hutchinson and Gigerenzer 2005). Though specific for a particular context rather than general purpose, these decision rules are not necessarily limited to a content domain. Adams et al. (this volume) provide behavioral and neural data on decision making in rhesus macaques (*Macaca mulatta*) which suggests that the same decision algorithms can apply across social and nonsocial domains. In particular, they argue that when macaques forage for food or search for information in their social environment, they use a comparable decision rule that simply weighs their return (on food or information) against a fixed threshold.

Work on foraging in humans has demonstrated that people use similar decision rules in very different types of foraging; for instance, physical foraging for food in an artificial fishing task and cognitive foraging for words in memory (Hutchinson et al. 2008; Wilke et al. 2009). Though applicable across content domains, these rules are by no means general purpose. Instead, they are likely tuned to the statistical structure of the environment (Hills et al. 2008). Thus, we might expect similar decision rules to operate in the foraging, social information search, and memory retrieval domains if all of these problems involve a common statistical problem to solve, such as finding objects in a clumped distribution.

Despite the advantages of specialized decision rules, disadvantages exist as well. An important difficulty faced by the toolbox approach is the *strategy selection problem* (Gigerenzer and Gaissmaier 2010; Marewski and Schooler 2011). As discussed by Kacelnik (this volume), Kurzban (this volume), and Mussweiler (this volume), having an adaptive toolbox requires selecting the correct tool for the current circumstance. How does an organism “know” which rule or strategy to use in which situation?

Kacelnik (this volume) criticizes the decision rules approach for ignoring the role of learning. He argues that, rather than selecting among specific decision rules, organisms can and do use reinforcement learning to solve many different problems. The question is: What is learned? Kacelnik assumes that the behaviors are what are learned. There is evidence, however, that strategies are learned as well. Humans, for instance, can learn to use the appropriate strategy for a given decision environment (Rieskamp and Otto 2006). Learning, therefore, leads to selecting the correct tool from the toolbox and thus a solution to the strategy selection problem. Nevertheless, more work is needed to address how organisms choose among the many possible decision strategies that they can implement. A Darwinian account could provide fruitful insight into this question.

A second (and related) critique of the decision rules approach is that we already have good evidence of a general-purpose decision-making mechanism, namely reinforcement learning (Kacelnik, this volume; Dayan, this volume). Though learning can apply across a range of circumstances, it is by no means unbounded. A Darwinian approach reminds us of the restrictions on learning. As an evolved mechanism, learning faces constraints and biases tailored toward evolutionarily relevant problems.

### **Learning Prepared by Evolution**

Since learning plays an extremely important role in human decision making, it is often thought that evolutionary biology cannot contribute much to the study of human behavior. In fact, many scholars in the humanities and social sciences consider decision making as a process governed mainly by experience and the imprints of culture in our minds (e.g., Sahlins 1976). However, experience needs to be acquired and cultural knowledge has to be gained. Evolution certainly has shaped the fundamental learning mechanisms by which this can be achieved. More generally, for any given animal species, the learning mechanisms are likely tailored in such a way to facilitate learning the specific things that matter for survival and reproduction under ecological conditions typical for their evolution. In other words, rather than being an omnipotent tool, learning is prepared to satisfy the particular (historic) needs of a species in an efficient way. The evolutionary preparedness of learning is thus fundamental to understanding mechanisms of decision making. It explains the impact of culture on humans’ choices of action as well as the difficulty a rat has in learning

to associate nausea with the ring of a bell (see Hammerstein and Boyd, this volume).

The specificity argument has its limitations in that evolution cannot easily respond to emerging needs of a species by developing appropriate mechanisms entirely from scratch. It usually has to modify existing machinery, and the modified machinery will thus be subject to “optimization under phylogenetic constraints.” Furthermore, there seems to be little use of loading the same brain with a great variety of different learning procedures, one for each problem. Even if almost unlimited mental resources allowed this to happen, an evolving species could then easily be trapped by “overfitting” its learning machinery to environments that are short-lived on the evolutionary timescale (Gluck et al., this volume). In the evolutionary picture of learning, there is thus space for psychological laws and principles that are valid under a wide range of conditions.

Such principles have been found in experimental psychology, but they percolated little through the disciplinary boundaries between psychology and biology. Kacelnik (this volume) rightly complains that, from its start, behavioral ecology addressed behavioral mechanisms as if experimental psychology never existed. In his view, established psychology—with its search for generality—was replaced by a search for rules of thumb which were generated and abandoned in an ad hoc fashion to interpret experiments within the narrow scope of their setting. Kacelnik considers one of the “flagships” in theoretical behavioral ecology, the marginal value theorem (Charnov 1976), and argues that predictions obtained from “sailing this ship” can be significantly improved and almost independently achieved with the findings of experimental psychology in mind.

Dayan (this volume) demonstrates nicely how the reasoning about general psychological principles can be combined with reasoning about specific adaptations. Looking at Pavlovian control of behavior, he stresses the broad range of problems that can be addressed through this mechanism. Dayan also emphasizes that animals have a rather limited repertoire of specific actions that evolution has “found to be useful” and which are triggered through Pavlovian learning. Pavlovian control thus seems to combine generality with specificity in a way that allows animals to cope quickly with variation in their typical environments.

Social animals are prepared to learn both individually and socially. Evolution tunes the balance between these mechanisms for acquiring information and made the human species uniquely dependent on social learning. Spreading rapidly all over the globe, humans needed quick responses to new environments which included the development of new tools and social arrangements. Hammerstein and Boyd (this volume) describe how learning from each other enables us to accumulate information across generations and acquire the tools, beliefs, and practices that single individuals could never have invented. Almost paradoxically, the accumulation of knowledge about adaptive practices hinges to a large extent on the fact that people often do not understand *why* culturally

transmitted behavior is adaptive. The crucial point here is that children learn to do what they are supposed to do without much cognitive interference. Human learning is biased toward conformism, and children receive most of their cultural information from older individuals who have a tendency to discourage questions by the young learner. This has the advantage that we do not waste our time trying to figure out what may be difficult or impossible to understand.

Do we really find it convincing that conformist elements in human learning govern our decisions toward adaptive behavior? Conformism means that we also learn to do things that have no adaptive value and which may even be harmful. Despite this drawback, evolution can favor forms of conformism strongly enough to induce occasional maladaptive “herding effects” in populations. A population then goes through phases where most, if not all, of its members make harmful decisions (Hammerstein and Boyd, this volume). This insight from evolutionary theory is of relevance to financial markets and helps us understand what economists call the “foolishness of the crowd.” Understanding the adaptive nature of and constraints on different types of learning can inform how learning is used in decision making.

### **Mental Adaptations to Sociality**

As discussed in the previous section, human learning is evolutionarily prepared to lead us through life in a highly social environment. Without this preparedness, particularly the conformist tendencies, human culture could hardly exist, and no one would have ever built a kayak, airplane, or spaceship. Human sociality is also based on a number of mental mechanisms beyond learning, which will be described below. These are superb cognitive features but not the ones on which economic decision theory is built. Economists typically envision a merchant who has learned arithmetic at school, effectively treating information processing as a “black box” (Bugnyar et al., this volume). In our evolutionary past, however, there was no teaching of mathematics and our Stone Age ancestors did not possess an abacus, slide ruler, or pocket calculator. They had to rely on means other than calculation to make their decisions. For this reason, and guided by empirical research on child development, Warneken and Rosati (this volume) advocate an alternative way of thinking about altruism and cooperation.

In Warneken and Rosati’s view, a decision to help or collaborate may be driven more strongly by intention attribution than by explicit calculations of costs and benefits. Children are able to differentiate intentions from behavioral outcomes and can tell, for example, whether a person is unwilling to hand them a toy or is unable to do so. This attribution of intentions makes it possible for the child to direct cooperative behavior preferentially to persons with cooperative intentions. The identification of intentions seems more important to social partner choice than engaging in the kind of payoff calculations around which economic theory is built. Chimpanzees (*Pan troglodytes*) share with us



some of the intention-attribution skills, yet we seem to outcompete them in our ability to form *joint intentions*—an important mental adaptation to sociality (Tomasello et al. 2005).

Emotions are another prominent feature of our minds, and sociality crucially depends on some of them. Conventional decision theory is surprisingly devoid of this issue and, as Jensen (this volume) puts it, philosophers have long viewed emotions as the enemy of reason. Jensen gives his picture of how the emotions govern positive and negative concerns for others. Like Warneken and Rosati, he emphasizes our ability to detect others' goals and describes positive social concern as an emotional state that motivates the actor to reduce the suffering of others and to seek their emotional well-being. Jensen disagrees with the idea of "psychological hedonism," in which altruistic behavior is seen as a selfish attempt to obtain internal rewards with no genuine concern for the recipient of the altruistic act. In his view, the primary objective of altruism is the well-being of others and not the internal reward that comes with it.

Not all social concern is positive, as suggested by the terms "moralistic aggression," "punitive sentiments," or "moral outrage." People (including children) do punish others for causing harm even if they have not been harmed themselves. Jensen (this volume) promotes the idea that human properties like spitefulness, *schadenfreude*, jealousy, and envy may be important enforcers of cooperation that stabilize prosocial behaviors and may be regarded as the backbone of human prosociality. According to Jensen, much less evidence exists for such a backbone in the social life of chimpanzees. Hammerstein and Boyd (this volume) emphasize that conventional decision theory has always made humans look intellectually superior to chimpanzees but it failed to grasp the important emotions to which Jensen discusses in his chapter. For example, game theorists found a flaw in the logic of deterrence and convinced themselves that it would never work in a world of rational decision makers. Their conclusion was that in real life, deterrence can only work because humans are not rational and also do not view their opponents as rational players.

Social psychologists and anthropologists, on the other hand, have actually studied a commitment device used in deterrence: anger (Nelissen and Zeelenberg 2009). The emotion anger prevents in-depth reasoning, causes us to take great risks, and sometimes carries us into absurdly costly forms of retaliation (e.g., road rage). Hammerstein and Boyd view emotions as both promoters and inhibitors of sociality and make the point that the targets of anger are shaped by cultural evolution. The degree of violence in response to transgressions, for example, may differ dramatically between two cultures if people value personal honor more strongly in one than in the other. Anger can, in principle, stabilize any norm defined by culture, not all of which are beneficial to society. Anthropologists have indeed described a variety of norms that are deleterious, such as mortuary cannibalism (Whitfield et al. 2008).

Mussweiler (this volume) argues more generally that the complexity of social information in our everyday life can only be handled through a number of



selective steps that determine (a) the information to which we should attend, (b) the information-processing tools to be used, and (c) the set of behavioral options from which one will finally be chosen. These three steps are ignored in conventional decision theory, where decision makers make use of all information they have, possess only one tool for processing it (utility maximization), and take all behavioral options into account. Evolution, however, has fabricated a fundamentally different decision maker who may look less ingenious to mathematicians but works astonishingly well in practice. To understand real decisions, we must understand how natural selection has shaped the selection steps highlighted by Mussweiler. Adams et al. (this volume) agree that social stimuli engage specialized mechanisms for the acquisition and processing of social information, but they assert that the decision rules used may be the same as those used in nonsocial situations. These perspectives highlight the need to explore the specificity of cognitive mechanisms to social situations. Nevertheless, they indicate that a “bottom-up” approach of investigating the cognitive mechanisms is critical to understanding the evolution of decision making in a social world (see also Bugnyar et al., this volume).

### **Error Management**

We have already seen that a Darwinian decision maker may be biased in (a) the contexts in which the decision mechanisms operate adaptively, (b) the types of information that can be learned easily, (c) the way that social information is filtered and processed, and (d) the triggers of emotional responses. Error management theory (Nettle, this volume) predicts that we will see biases in the types of errors made when we make inferences about the state of the world (Haselton and Buss 2000). Not all errors are created equally.

From the evolutionary perspective, the handling of errors in the decision-making machinery must reflect the effect of these errors on an organism’s survival and reproduction (fitness). When different types of errors result in different costs, we would expect to see biased decision making. Natural selection will likely favor the avoidance of even small errors if they incur high costs in terms of fitness. In contrast, seemingly large errors (e.g., a male mating with a member of the wrong species) may not face strong selective pressure if they have little impact on fitness. This is referred to as the “smoke detector principle” (Nesse 2001b). A smoke detector’s job is not to minimize the numbers of errors it makes. Its job is to detect the presence of smoke in order to save lives. In doing its job, smoke detectors are biased to give an alarm when no smoke is present (a false alarm) rather than fail to give an alarm when smoke is present (a miss). Many inattentive cooks have suffered the indignity of fanning at a smoke detector after burning a meal. Despite the inconvenience and embarrassment of a false alarm, fanning a smoke detector or even evacuating a building is preferable to a smoke detector not triggering in the presence of a

real fire. We will pay the minimal costs of false alarms to avoid the devastation of a miss.

Kareev (this volume) argues that this bias toward avoiding misses is a core property of human decision making based on the potential benefits of using our limited short-term memory to detect patterns in our environment. This is an appropriate strategy when the costs of misses outweigh the costs of false alarms, as is seen in the smoke detector example, as well as when avoiding predators, predicting the impact of a looming object, attributing agency to objects in the environment, and detecting signals of sexual interest from potential mates (Nettle, this volume). When false alarm costs outweigh those of misses, however, we see the opposite pattern with greater sensitivity to false alarms. For instance, females may be biased toward accurately detecting honest signals of male parental investment (Haselton and Buss 2000). Missing an investing male is not as costly as succumbing to the false advertising of a deadbeat.

Considering the costs of errors, therefore, is critical to understanding why we observe biases in decision making. Though signal detection theory and expected utility theory also incorporate the costs of errors, Nettle (this volume) argues that the evolutionary approach via error management theory makes key predictions about the kinds of contexts in which we would expect to see biased decision making; namely, contexts with important implications for evolutionary fitness. Moreover, due to the lag in natural selection's ability to adapt organisms to their environment, error management theory can explain potential biases in situations in which no current cost differential exists, though historically strong evolutionary pressures may have resulted in divergent costs. Thus, while the occurrence of errors as such relates to mechanistic properties of the mental machinery, the management of these errors cannot be understood without exploring the evolutionary question of which errors in decision making are tolerable and which are not.

### **Robustness and the Mechanisms behind It**

From an engineer's point of view, robustness is the ability of a system to maintain its functionality across a wide range of operational conditions. Different conditions arise, for example, from environmental variation, noisy input, sloppiness or breakdown of system components, and subversion by parasites. In the course of evolution, organisms are expected to adapt their behavior to the variety of conditions under which they have to survive and reproduce. Robustness as a concept is, therefore, of great relevance to the Darwinian approach to decision making. Flack et al. (this volume) describe the various ways in which biologists have used this concept at different levels of organization: from molecular systems to individual decision making and animal societies.

Most organisms operate in a highly variable and complex world. How do organisms make decisions when facing large temporal variation and spatial heterogeneity? To meet the enormous challenges, humans and other species

must detect regularities in their new physical and social environments. At first glance, our brain seems poorly equipped for this task. Our short-term memory, for example, can hold just a few items at a time. This may seem ridiculously small for a storing device. Kareev (this volume) makes the interesting point, however, that limited short-term memory actually has a number of advantages when individuals are in search of unknown regularities. It forces us to do much of our “mental statistics” on the basis of small samples. Correlations are then likely to appear stronger than they actually are and variance is typically underestimated. This amplification of correlations makes it easier to detect the regularities and, in this sense, improves our mechanisms for exploration. There is a drawback, of course, since occasionally we will find correlations that do not exist at all (i.e., the world will look more regular than it is). But, returning to error management theory, it may be more costly for individuals to overlook important regularities than to “fantasize” a few.

Kareev (this volume) offers further arguments why it can pay to rely only on a very limited number of recent experiences in decision making. Such a self-imposed restriction can make it easier for organisms to follow changes in the environment because it helps them avoid overfitting to conditions that vary in space and time (Gluck et al., this volume). Furthermore, remembering too many past events involves the risk of always lagging far behind in fast-changing environments. Dayan (this volume) emphasizes the fact that information stored in our working memory degrades more or less gracefully as memory is taxed. The evolved working memory’s design thus differs considerably from that of a human-designed computer, which stores information with extreme reliability in addressable locations until a voluntary act of deletion occurs.

Gallistel (this volume) has difficulties accepting the idea that the brain is fundamentally different from a computer in that it lacks an addressable read-write memory. He explains why such a memory is taken by computer scientists to be the foundation of any powerful computing machine. He then reviews evidence that behavior is mediated by computational information processing that deals with extensive data structures, as demonstrated, for example, by experiments that reveal the contents of the cache memory of food-caching birds (Clayton et al. 2001a). He asks whether the conceptual chasm between the computer science understanding of the essential role that a symbolic, read-write, addressable memory mechanism plays in any powerful computing machine and the neuroscience conception of memory, which is not symbolic, not read-write, and not addressable, is a problem for computer science or a problem for neuroscience.

Among the obstacles in addressing environments are the inherent uncertainties that need to be handled, which arise from our ignorance as well as from environmental change. Dayan (this volume) discusses how evolution has prepared our brains for the challenges posed by uncertainty. For example, in his view the brain “offers” itself an exploration bonus when assessing the payoffs relating to actions. This bonus supposedly drives our exploratory behavior.

But how do we know about the existence of such a quantity? Dayan interprets some experimental findings from neuroscience as keys to how the bonus manifests itself in the brain. Novel objects, for example, generate temporary activity in the dopamine system that resembles the activity triggered by unpredicted “true” rewards.

Discussing robustness at a higher level of abstraction, Dayan (this volume) sees two sources of noise that pose a major threat to it. There is noise associated with incomplete and inefficient learning in what he calls the model-free system, and there is noise inherent in the complex calculations performed by the brain with its limited potential for computation. The latter limitation becomes particularly visible when our decisions are based on internal models of reality. These models enable us, however, to predict events under changed conditions long before learning could achieve anything.

Kurzban (this volume) takes a perspective on robustness in which the brain is already equipped with a number of tools that allow us to act as if we had indeed modeled some aspects of reality. He argues, for example, that our ancestors typically encountered a spatially and temporally autocorrelated world when searching for water, food, and other important items. The a priori expectation of autocorrelation may thus be one of the innate biases that evolution has implemented in our decision-making machinery. This would explain why human predictions are often based on the implicit assumption that events come in “streaks” and are particularly likely to occur after they just occurred—the “hot-hand” phenomenon (Wilke and Barrett 2009).

A final word must be said about redundancy, an extremely important design principle used by human engineers in their efforts to create robust machinery (Flack et al., this volume). Space ships are, for example, equipped with several computers, each of which performs the same calculations. If one of them makes a calculation error, this computer can be “outvoted” by the other two machines. Dayan (this volume) argues that the robustness of animal and human decision making is also supported by a fundamentally different kind of redundancy. Instead of just having duplicates of subroutines or other mechanisms, we have different computational devices that rely on very different procedures to perform similar tasks. This kind of redundancy protects against both errors in computing and maladaptive properties of the implicit models on which these computations are based. Dayan views decision making as a permanent struggle for the “brain’s attention” by different mechanisms that provide the kind of redundancy just described. Evolution seems to have equipped our minds with elements of internal competition to maintain robust decision making.

### **Biological Roots of Variation**

If there is one thing that remains constant in decision making, it is variability. Despite the sincere wish of economists—that a single equation and set of parameters can hold for all decision makers in all situations—this is not to be. For

instance, many researchers are interested in the rate of discounting that people employ (i.e., the rate at which future benefits are devalued). Frederick et al. (2002) contend that measures of discounting have not converged on a single discount rate as has happened with the speed of light. In fact, studies measuring discount rates yield values ranging from a negative discount rate (meaning a preference for delayed rewards over immediate rewards) to an infinite discount rate (meaning the strongest possible preference for immediate rewards over delayed rewards). Economists tend to sweep variation like this under the rug or treat it as noise.

Can a Darwinian approach account for variation? Though variation is a key component of natural selection, evolutionary game theory models typically predict a single best solution or small set of best solutions to an adaptive problem. We observe individual differences in behavior across a broad spectrum of species and in a wide range of contexts. In addition, within individuals we see frequent variation in behavior. Both between- and within-individual variation greatly exceeds that expected by many of these models, and understanding the role of this variation will be important in developing a theory of decision making.

To this end, biologists have turned to psychology as a field that takes variation seriously. To minimize anthropomorphism, some biologists refer to individual differences in the behavior of animals as “behavioral syndromes.” Bolder researchers use the same term as applied to humans: personalities. Evolutionary models are beginning to explore adaptive accounts of the breadth of variation between individuals in their behavior (Dingemanse and Wolf, this volume; Dall et al. this volume). However, these models tend to focus on behavioral polymorphism and do not account for many interesting aspects observed in individual differences.

A key attribute of personality is not just consistent individual differences but also stability across contexts. For instance, one of the classic personality traits in animals is the bold-shy continuum, in which individuals range from novelty seeking to neophobic. Dingemanse and Wolf (this volume) describe early work on stickleback, in which fish who act aggressively toward conspecifics intruding on their territory are also more likely to approach potential predators. This correlation of boldness appears across a range of species, and Dingemanse and Wolf (this volume) highlight a growing interest in the genetic and physiological mechanisms underlying similar behavior across contexts. For instance, researchers are conducting quantitative genetic analyses which show strong genetic components, molecular genetic analyses that reveal candidate genes, neuroendocrine analyses of coping styles and stress response, and correlations of metabolic rates with behavioral differences.

In addition to mechanisms underlying personality, biologists are now beginning to ask critical questions about possible advantages of both behavioral consistency and correlated traits (Dingemanse and Wolf, this volume). Do correlated traits provide benefits to individuals or are they simply by-products of the genetic and physiological mechanisms underlying behavior? Research on

spiders shows that, though bold individuals gain benefits in foraging situations, their aggressiveness imposes costs on fertility when aggressive females cannibalize potential mates before copulation (Johnson and Sih 2005). Thus, the correlated trait results in adaptive trade-offs. Would an individual be better off bold with food and shy with mates? Regardless of the adaptive benefits of correlated traits, they are real and therefore require our attention. Dall et al. (this volume) emphasize that this reality means that we cannot necessarily treat different adaptive problems as independent. As the previous example illustrates, the mating game is not independent from the predator avoidance game, and this has critical implications for how we model behavior. Understanding the adaptive value of and constraints on behavioral consistency and correlations will offer key insight into the evolution of decision making.

Individual differences provide one type of variability, but we also observe variation within individuals. Brown et al. (this volume), suggest that within-individual variation can result from noise, context, mood, life span changes, and prior experience. Noise in behavior can provide benefits when individuals want to exhibit unpredictable behavior. This can occur when trying to avoid predators or to make a credible threat of irrational behavior. Though some decisions appear rather capricious, a careful examination of the situation may highlight important context-specific predictors of behavior. Facebook notwithstanding, we often behave differently among family, friends, and strangers. Similarly, members of other species decide differently with a (potential) mate, rival, or dominant present. Even in the same context, our moods can have critical influences our actions. Brown et al. (this volume) describe how positive moods and pessimism shape decision making in humans and other animals. For instance, inducing a positive mood shifts people's risky decision making away from focusing on probabilities of reward and toward focusing on outcomes (Nygren et al. 1996).

Variation in decision making also occurs over the life span. Early in life, juvenile animals and human children demonstrate changes in decision making as various cognitive capacities come online during development (Jensen, this volume; Warneken and Rosati, this volume). At the other end of the continuum, we see modifications in decision making as adults age. Brown et al. (this volume) review research which suggests that cognitive aging results in increases in risk avoidance, patience for future rewards, weighting of losses compared to gains, and altruistic behavior. They also highlight the role of prior experience in decision making. The decision-by-sampling approach accounts for how sampling from the distribution of prior experience via long-term memory can influence decisions (Stewart et al. 2006). For instance, risky choices for mortality-related decisions correlate with the risk of mortality faced in a given participant's country (Olivola and Sagara 2009), suggesting that risky choices match the distribution of experience in the world. Thus, individuals with different experiences, or more specifically different memories, will exhibit different choices.

The study of variation from a biological approach remains in its infancy. At the moment, psychology has a lot to offer to an evolutionary account of variation in behavior both between and within individuals. Despite the standard tendency to “control for” this variation, the psychology of individual differences will be critical in informing a Darwinian approach to decision making.

### **Concluding Remarks**

After describing several reasons for invoking evolution in decision theory, we now discuss at a more general level to what extent this helps us understand the mechanisms of decision making. Let us first play the devil’s advocate and question the importance of marrying evolution with cognitive science. Most of the advancements in behavioral ecology, for example, were made by explicitly ignoring cognition and treating evolution as if natural selection acted directly on behavioral traits (the so-called “behavioral gambit”; Fawcett et al. 2012). Research programs in this field were particularly successful because they used the shortcut of circumventing the nitty-gritty of cognitive machinery. We must admit that using this shortcut has its rationale. Evolutionary theory cannot predict mechanisms as such, since in principle many different mechanisms can serve as a tool for solving the same problem. To use an analogy from engineering, there are many ways to design a clock, but all that a well-engineered clock can tell us with high accuracy is what time of the day it is.

As convincing as this multiplicity argument may sound in defense of classical research programs, it is also quite misleading. We use our wristwatch at different temperatures, for example, and may even leave it on while swimming. The mechanism operating the watch must therefore tolerate changes in temperature and pressure and continue to work adequately while being submerged in water. More generally, the more we know about the conditions under which a mechanism has to operate, the better we are able to reflect necessary specifications for its design. So, despite the fact that evolutionary theory cannot simply predict the entire mechanisms of decision making, it can inform us about fundamental properties that the evolved mechanisms can be expected to possess. As the chapters of this volume document nicely, these properties include (a) biases in learning, error management, and information usage, (b) robust responses to variation in the environment, (c) variation within and between decision makers, and (d) specializations for coping with complex social situations.

It is also important to state properties that are unlikely to exist. For example, one would not expect consistency to be a general property of evolved decision mechanisms (see, however, Chater as well as Hagen et al., both this volume). An evolutionary theory of decision mechanisms, therefore, strongly undermines the approach that dominated decision theory in economics for more than the last hundred years. Research combining evolution and cognition



does indeed give us good reasons to knock economic decision theory off its pedestal.

There are two more important aspects of evolutionary analysis which deserve general attention. One is the phylogenetic approach in which ancestral mechanisms are considered as starting points from which decision mechanisms evolved. Studying decision making in phylogenetically closely related species such as chimpanzees (Jensen, this volume; Warneken and Rosati, this volume) can provide unique insights into the human condition. Humans and chimpanzees differ, for example, in how they learn from others. However, chimpanzees do show sophisticated forms of social cognition that offer a foundation for understanding our own social decision making. The second important aspect of evolutionary analysis is the comparative approach; that is the study of how different species cope with similar kinds of problems. The comparative approach offers a glimpse into how the environment shapes decision mechanisms by comparing decisions across species that both face similar environmental pressures and are adapted to different environments. For instance, from honeybees to hummingbirds and locusts to starlings, we see similar effects of context on decision making that violate classical decision theory (Rosati and Stevens 2009; Hagen et al., this volume). The ubiquity of these behaviors across species suggests that natural selection has shaped decision mechanisms to solve a widely applicable problem. On the other hand, differences in environmental pressures can be used to predict differences in decision making, such as when foraging ecology matches temporal preferences in chimpanzees and bonobos (Rosati et al. 2007).

In summary, evolutionary theory is, of course, far from being anything like an omnipotent explanatory device. It can shed light on why all sorts of biases and apparently odd effects exist in human decision making, why *Homo sapiens* is far from being anything like a “relative” of *Homo economicus*, and why we are nevertheless quite successful in addressing our everyday problems. From an evolutionary perspective there seems to be a logic behind decision making in humans and animals, but it is a logic that makes individuals successful in real life without caring about axioms of rationality. We thus see the contours of a new decision theory and wish to merge cognition and evolution further in order to root this theory firmly in empirical grounds and make sense of the facts.

### Acknowledgments

On behalf of all of the participants, we wish to thank the Ernst Strüngmann Forum for their generous funding of this Forum and resulting publication. We are also grateful to Gerd Gigerenzer, who encouraged us to pursue this theme, and to the Santa Fe Institute, where some of the planning took place.

We are particularly indebted to the Forum’s program director, Julia Lupp, and her staff, Marina Turner and Andrea Schoepski, who managed everything in a highly professional and inspiring way. Their admirable personal engagement enabled us to focus

entirely on the scientific subject matter and enjoy the chairing of this Forum without having to endure any of the headaches that often accompany the necessary organization. Julia, Marina, and Andrea impressed us deeply with both their warmhearted and firm way of handling us scientists, who can be difficult at times.

The Forum also benefited greatly from contributions made by Nick Chater, John McNamara, and Reinhard Selten as members of the program advisory committee, as well as from the support received by the moderators of the four groups. Furthermore, this volume would not have come into existence without the brave efforts of the four rapporteurs—Ed Hagen, Kevin Gluck, Sasha Dall, and Thomas Bugnyar—who had to pull many strings together as they created the first drafts of group reports during a single afternoon (and long night) of the Forum while the rest of us were able to enjoy the sites and culinary delights of Frankfurt. Needless to say, we appreciate these efforts and are grateful to the help offered from everyone during the revisions of the chapters for this book.