Cell PRESS

Focus Issue: Overconfidence and deception in behaviour

The evolution of error: error management, cognitive constraints, and adaptive decision-making biases

Dominic D.P. Johnson[1], Daniel T. Blumstein[2], James H. Fowler[3], and Martie G. Haselton[4]

[1] Department of Politics and International Relations, University of Oxford, Oxford, UK
[2] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA
[3] Political Science Department and Medical Genetics Division, University of California, San Diego, CA, USA
[4] Departments of Psychology and Communication Studies and the Institute for Society and Genetics, University of California, Los Angeles, CA, USA

Counterintuitively, biases in behavior or cognition can improve decision making. Under conditions of uncertainty and asymmetric costs of 'false-positive' and 'false-negative' errors, biases can lead to mistakes in one direction but – in so doing – steer us away from more costly mistakes in the other direction. For example, we sometimes think sticks are snakes (which is harmless), but rarely that snakes are sticks (which can be deadly). We suggest that 'error management' biases: (i) have been independently identified by multiple interdisciplinary studies, suggesting the phenomenon is robust across domains, disciplines, and methodologies; (ii) represent a general feature of life, with common sources of variation; and (iii) offer an explanation, in error management theory (EMT), for the evolution of cognitive biases as the best way to manage errors under cognitive and evolutionary constraints.

The ubiquity of error

All of us face the problem of balancing the potential costs of alternative decisions in everyday life. Should you risk a detour to avoid traffic? Should you bother paying for travel insurance? Should you put your savings in the bank or the stock market? Such dilemmas occur whenever decisions are made under uncertainty and alternative outcomes are likely to incur different costs – a common scenario in which many of our decisions turn out to be wrong. The problem is not new. Humans and other organisms have had to deal with balancing risks for millions of years [1–6] and there is evidence that we have evolved specific *biases* that help to minimize the costs of mistakes over time. Recent work in a range of disciplines invokes the same logic and we suggest that all such decision-making problems fall under the unifying theoretical framework of error management.

Error management is important because how people tend to balance alternative possible errors, as well as how they *should* balance these errors, has consequences for a range of challenges of the 21st century, spanning medicine (e.g., whether and how to treat cancer), public policy (e.g., how to invest pension funds), engineering (e.g., what magnitude of earthquakes nuclear facilities should be built to withstand), international security (e.g., when to act against states that might be developing nuclear weapons), and climate change (e.g., how much to invest in or how to enforce carbon emission reductions). In 2000, Swets *et al.* lamented that error management techniques were 'virtually unknown and unused' in many fields, despite offering an extremely valuable way of improving decision making [7].

In the following sections, we: (i) outline the logic of biased decisions as adaptive; (ii) introduce EMT; (iii) explain why humans would have needed to evolve biases to achieve adaptive behavior; (iv) present a range of interdisciplinary studies identifying the phenomenon of error management; (v) consider error management's scope as a unifying framework for understanding a wide range of judgment and decision-making problems; and (vi) offer an explanation for why there have been fundamental differences of opinion among scholars on the evolution of biases.

Biased decisions as adaptive

In recent decades, economists and psychologists have documented a long list of biases in human judgment and decision making [8–11], with important consequences for economics, politics, and society [12–14]. Rather than being mere quirks of human nature, however, there is growing evidence that these biases represent adaptive solutions to the decision-making problems of our evolutionary past [15–18]. Evolutionary biology therefore has an important role to play in understanding decision-making biases, especially given that their underlying logic has long been examined by ethologists and ecologists [4,5,19–22]. For example, when resources are plentiful and dangers scarce, organisms should avoid risky decisions, just as standard economic models would predict [23]. However, when starvation or other dangers threaten survival (and thus expected reproductive value, in the eyes of evolution), selection can favor

Corresponding author: Haselton, M.G. (haselton@ucla.edu).

0169-5347/$ – see front matter
 http://dx.doi.org/10.1016/j.tree.2013.05.014

Table 1. The four possible combinations of beliefs about the world (X) and actual states of the world (giving rise to two different types of error)

		Actual	
		X	Not X
Belief	X	Correct	False-positive error
	Not X	False-negative error	Correct

organisms (or strategies) that take whatever risks might be necessary to give them a chance of life rather than death. These strategies do not necessarily maximize *expected payoffs* (of, say, food). Instead, they maximize *Darwinian fitness*, which can be a nonlinear function of expected payoffs [4]. Hence, judgment and decision-making 'biases' (as judged by comparison with standard economic models) can be expected and appear to be primarily geared toward the management of alternative errors.

What is Error Management Theory (EMT)?

A decision can be wrong in two ways (Table 1). It can assume X when not-X is true (a false positive) or not-X when X is true (a false negative). The principle of error management suggests that, under uncertainty (that is, where the true probability of outcomes cannot be precisely predicted), and if the costs of false-positive and false-negative errors are different, an effective decision-making strategy can be a *bias* toward making the least costly error over time (a generalized illustration is given in Figure 1). The conditions for error management might seem limiting, but are common to a wide range of decision problems; the true probability of outcomes is rarely known precisely and different outcomes are unlikely to have identical costs.

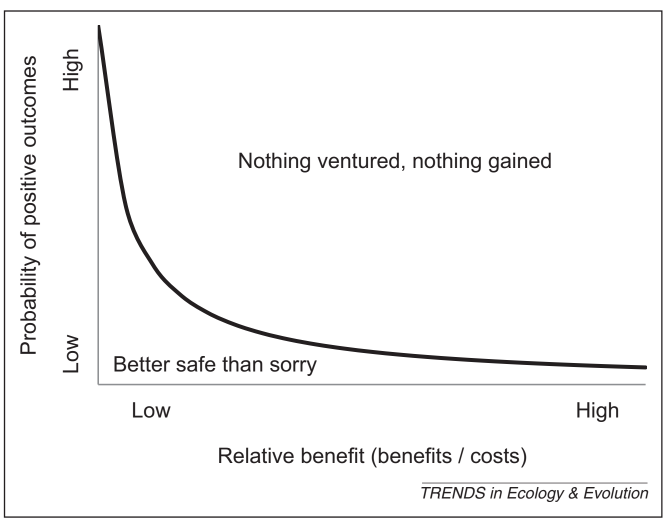

Figure 1. A generalized illustration of error management showing how changes in the probability (y-axis) or relative benefits (x-axis) of outcomes can affect biases. The curve depicts the probability of success of some decision over a range of benefits and costs. Where benefits exceed costs, *overestimating* the true probability constitutes an assessment error but nevertheless leads to the 'right' decision (e.g., to decide to act when the odds are favorable). By contrast, *underestimating* the true probability also constitutes an assessment error, but this time leads to the 'wrong' decision (e.g., to decide not to act despite the odds being favorable). This logic is reversed where costs exceed benefits. In general, when the probability of positive outcomes and/or the relative benefits are low, risky behavior should decrease and actors should engage in more conservative 'better safe than sorry' behavior. As the probability of positive outcomes and/or the relative benefits increase, risky behavior should increase and actors should engage in bolder 'nothing ventured, nothing gained' behavior.

Error management as an general engineering principle

Smoke alarms provide an illustrative example of error management as a general principle of engineering [24]. Smoke alarms are deliberately set to go off what might seem to be 'too often' (sometimes they go off when you burn your toast), but this is only to avoid making a much worse error in the other direction (failing to go off when there is a real fire). Because there is uncertainty in the signal (the aerial particulates might appear similar to the smoke alarm in the two cases), the only way to make sure all genuine fires are detected is to set alarms to be extremely sensitive – we give them a bias. We might think the alarm has been set to be 'too sensitive', but they are set to go off just the right amount, given the available data on the dangers of real fires.

Error management as an evolved heuristic to exploit the engineering principle

Error management theory (EMT) attracted particular attention (and the term was coined) when it was applied to explain decision-making biases in *humans*. If decision-making errors had asymmetric costs over human evolutionary history, natural selection might have favored cognitive biases that minimized whichever mistakes incurred the greatest costs [15,24,25]. This intriguing logic led to a number of studies that proposed a role for EMT in explaining puzzling empirical biases such as sex differences in dating behavior [26,27], overoptimism [28], cooperation toward strangers [29], and belief in supernatural agents [30].

Note that although EMT focuses on costs, benefits are not ignored. It is the net payoffs – costs plus benefits – of false-negative and false-positive decisions that matter. EMT predicts a bias in whichever direction maximizes fitness. This tends to be worded in the EMT literature as 'minimizing costs' and this is a legacy of the original puzzle: how can we explain apparently costly behaviors in adaptive terms?

Why do we need biases to achieve adaptive behavior?

If we are aiming to assess the true probability of some event (e.g., whether a car is coming when we cross the road), but there is uncertainty in this probability, mistakes can occur in both directions – that is, sometimes assuming a car is coming when one is not or sometimes assuming no cars are coming when one is. A potentially efficient way to avoid the latter, much more catastrophic error is to develop a bias to act *as if* a car is coming whenever we cross the road [31]. Recent work shows that the asymmetry of costs does not have to be large to favor a bias [32,33]. Although this example suggests why a bias can be beneficial, it is insufficient to identify the *cause* of such a bias.

Adaptive behavioral biases

If one could accurately estimate the costs and probabilities of events, one could manage errors through 'cautious-action policies' [31] (one form of which is Bayesian updating [22,32]). For example, we might develop a behavior to look before crossing the street not because we have an in-built fear of cars but because we calculate, learn, or discover that the costs of failing to do so are severe and we adjust our behavior accordingly – investing considerable time to look even when no cars are close. To a third party, this *behavioral bias* may appear to be 'economically irrational' because we are consistently wasting time worrying about an event that rarely happens, but if it helps us to maximize our chance of survival (and reproduction), it is 'adaptively rational'. In other words, our behavior is not tuned to fit a model of expected *payoffs*, but a model of expected *fitness*. For this reason, a behavior that seems 'biased' would not seem biased if we had complete information about the consequences of those actions for survival and reproduction. Short-term costs can have long-term gains and may be measured in different currencies.

Adaptive cognitive biases

To understand why we need a *cognitive bias* – one that skews our assessments away from an objective perception of information – we need to pay attention to cognitive constraints (the limitations of brains as decision-making machines) and evolutionary constraints (natural selection as an imperfect designer).

Decisions involve a complex array of variables and we are often unable to keep track of accurate statistical information about the probabilities and costs of false-positive and false-negative errors (we might not even correctly perceive errors, let alone their future fitness consequences, when they occur). Even if we did have such information, the processing of multiple relevant variables would be cognitively taxing and itself subject to error. Natural selection, therefore, might have favored biases that circumvent the complex calculations otherwise needed for optimal decisions (rather than human brains having to do these calculations on the spot) [34,35].

Evolutionary constraints suggest additional reasons why a cognitive bias might have been favored by natural selection rather than relying on cautious-action policies [33,36]: (i) efficiency – a cognitive bias might have been biologically cheaper to produce or operate; (ii) speed – a cognitive bias might have allowed faster responses; (iii) evolvability – a cognitive bias might have been more readily available due to pre-existing cognitive machinery; and (iv) the adaptive landscape – a cognitive bias might have offered a way to climb a local fitness peak (maximizing fitness given immediate options), even if a better solution lay across a fitness valley that natural selection could not cross. Such valleys cannot be 'crossed' because a population would have to systematically endure fitness damage on the way, which selection would halt.

In short, cognitive biases are likely to have evolved because there are biological limitations to decision making: namely information, processing, efficiency, speed, evolvability, and the adaptive landscape. New work in neuroscience supports the notion that biases are often cognitive and not merely behavioral [37,38]. Empirically, humans do not seem to live up to the commonly vaunted ideal of optimal Bayesian decision makers [39]. As Nobel laureates Daniel Kahneman and Amos Tversky concluded in response to the idea that people might be at least approximately Bayesian, 'man is apparently not a conservative Bayesian: he is not Bayesian at all' [40].

Interdisciplinary convergence on error management

In this section, we present a series of examples of error management (with more in Table 2). EMT was formulated by evolutionary psychologists to explain how apparently costly psychological biases might in fact be adaptive [15,24–26,28,41]. However, we have found a wide range of literature from different disciplines and domains in which the same logic has been invoked. In essence, many scholars have independently discovered the logic of error management, demonstrating its robustness across widely different methodological approaches, modelling techniques, and domains of application. Some examples occur in biological *systems* (reflecting error management as a general engineering principle), such as the evolution of allergies as 'the price we pay' for protection against parasites [24,42], preferential protection or repair to reduce mutation rates in critical parts of the genome even if this incurs a cost of lower protection for other genes [43], or a bias in thermoregulation to manage asymmetric fitness costs of temperature increases and decreases [44]. In the longer examples outlined below, however, we focus on instances of error management in the minds or behavior of *individuals* (reflecting evolved biases that *exploit* the engineering principle of error management).

Varieties of error management

An early application of error management comes from animal behavior. Ethologists coined the 'life–dinner principle' to explain why it is better to err on the side of caution when foraging around predators (it is better to miss a meal than lose your life) [45]. This was generalized into a large literature on balancing the risks of alternative decisions, most notably in optimal foraging theory (for reviews, see [19,46,47]). For example, Bouskila and Blumstein [48] showed that predation-hazard assessments, which by their nature are imperfect, can become biased. They found a 'zone of tolerance' in which decisions should be biased toward overestimating predation risk. In this region, overestimation was marginally costly to expected fitness but less so than the potential costs of predation. As would be expected, the degree and direction of such a bias varies with context (Box 1) [49]. In short, these studies identified error management – in this case, a bias in the assessment of foraging and predation risk.

In a different area of animal behavior, biologist R. Haven Wiley applied signal-detection theory to animal communication [50]. He identified the matrix of false-positive and false-negative errors and argued that animals commonly make mistakes in the service of attempting to avoid bigger mistakes. Uncertainty was evident in the fact that signals are hard to discriminate, especially at long range, in aggregations of multiple individuals and where there is deception. Depending on ecological parameters, the differential costs of errors can lead to 'adaptive gullibility' (susceptibility to deception) or 'adaptive fastidiousness' (low responsiveness to signals). The latter scenario may in turn lead to the evolution of exaggerated signals. Again, we have an example of error management – in this case, a bias to over- or under-value signals.

Psychologist Ryan McKay and philosopher Daniel Dennett published a target article in *Behavioral and Brain*

Table 2. Studies invoking error management

Study	Discipline	Focus	Analytic strategy	Conclusion
Pascal (1669) [60]	Theology	Belief in God	Narrative logic	'Pascal's wager': the costs of atheism and benefits of belief in God are so great that one should assume God exists ('if you gain, you gain all; if you lose, you lose nothing')
Green and Swets (1966); Swets (1992) [7,61]	Psychophysics	Decision thresholds (e.g., in defense, aviation, and medicine)	Mathematical model	Decision thresholds depend on the relative costs of errors, benefits of correct decisions, and prior probabilities of events
Egan 1975 [62]	Psychophysics	Signaling theory	Mathematics and experiments	A response bias can in some circumstances be a better way of achieving decision goals
Bouskila and Blumstein (1992) [48]	Animal behavior	Antipredator behavior in animals	Mathematical model	Animals who overestimate predation risk have reduced mortality relative to unbiased or underestimating animals
Wiley (1994) [50]	Animal behavior	Animal signaling systems	Mathematical model	'Adaptive gullibility' (stable strategy of being easily deceived) and 'adaptive fastidiousness' (stable strategy of low responsiveness to signals) evolve depending on the relative costs of errors and benefits of correct decisions
Cosmides and Tooby (1994) [35]	Evolutionary psychology	Biases in economic behavior	Narrative logic	Specialized information processors ('biased' with information supplied by natural selection) outperform general-purpose, content-general 'rational' models of mind
Higgens (1997) [63]	Social psychology	Biases in human psychology	Narrative logic	Whether an individual is biased toward pursuing pleasure or avoiding pain depends on psychological needs and ideals
Tetlock (1998) [56]	International relations	Perceptions of enemy intentions	Narrative logic	States that overestimate the probability of other states having aggressive intentions (and arming in response) make a better error than assuming they are benign and risking deterrence failure, exploitation, or attack
Pacala *et al.* (2003) [64]	Environmental conservation	Estimates of environmental harm from human causes	Informal quantitative model	'Environmental alarms' set too low because the cost of inaction is huge relative to the cost of action
Haselton and Buss (2000) [26]	Evolutionary psychology	Biases in human psychology	Narrative logic and experiments	In courtship communication, selection favored an overestimation bias in men's inferences of women's sexual interest and an underestimation bias in women's inferences of men's intentions to commit to a long-term relationship
Nesse (2005) [24]	Evolutionary medicine	Immune and other body defenses	Mathematical model	False alarms expected and tolerated in defense against harms with great costs
Haselton and Nettle (2006) [15]	Evolutionary psychology	Biases in human psychology	Mathematical model	'Optimistic' and 'pessimistic' biases differ by domain depending on the relative costs of false-positive and false-negative errors over evolutionary history
Foster and Kokko (2009) [53]	Evolutionary biology	Superstitious behavior in humans and animals	Mathematical model	Selection favors strategies that lead to frequent errors in associations of cause and effect as long as the occasional correct response carries a large fitness benefit
Johnson (2009) [30]	Evolutionary biology	Belief in supernatural agents	Narrative logic	Individuals who believe their actions are observable and punishable by supernatural agents might be more likely to avoid the costs of social transgressions
McKay and Dennett (2009) [31]	Philosophy	False beliefs in human psychology	Narrative logic	True beliefs most often adaptive, but 'positive illusions' sometimes have a fitness advantage over true beliefs
McKay and Efferson (2010) [32]	Economics and psychology	Biases in human psychology	Mathematical model	Systematic departures from Bayesian beliefs can be adaptive when Bayesian updating is constrained
Johnson and Fowler (2011) [33]	Evolutionary biology	Conditions under which overconfidence can evolve	Mathematical model	Overconfidence is an evolutionarily stable strategy as long as benefits are sufficiently large compared with costs, and there is uncertainty in the assessment of opponents

Sciences reviewing the evidence for adaptive advantages of a wide range of 'false beliefs'. They identified EMT as a key potential driver of the evolution of such biases [31]. In particular, they highlighted 'positive illusions' (overestimating one's qualities, control over events, and future) [51,52] as false beliefs that plausibly confer adaptive advantages. This offers another example of error management – a bias that promotes success by increasing ambition and perseverance.

Around the same time, evolutionary biologists Kevin Foster and Hannah Kokko presented a model to explain 'superstitious' behavior, in which animals overestimate the

Box 1. Sources of variation in managing errors

Temporal constraints

Risk taking increases when time is running out and fewer opportunities are available [65]. Consider the 'Hail-Mary' football pass as the clock ticks down, the relaxation of mate choice criteria at the end of the mating season or lifespan, or the approach of closing time in a bar [66,67]. The optimal level of bias also depends on aspects of life history such as strength, status, age, and hunger [68]. For example, small birds increase risky foraging behavior as the end of the day approaches, as securing enough energy to get through the night becomes critical [69].

Ecological constraints

Risk taking increases when options are constrained by context. For example, as population density increases, energy reserves among individuals drop and the probability of fighting over resources – a risky strategy – increases [70]. Indeed, models suggest that lethal fighting can evolve if the stakes are high enough [71]. Similarly, optimal foraging becomes more risky under certain food distributions [20]. In general, environmental factors affect the emergence and direction of bias [21,49].

Costs

Risk taking increases if the relative costs of false-positive and false-negative errors change. Risks are more worthwhile if the prize is greater. Humans appear to be particularly susceptible to these kinds of change, because we readily envisage benefits (such as a $10 million prize), but not probabilities (such as a 1 in 10 million chance of winning it). Thus, more people enter the lottery when the prize is higher, though the probability of winning remains infinitesimal (and actually declines because more people are competing for it).

Probabilities

Risk taking increases if the probability of success increases. The total cost of false-positive or false-negative errors is actually the *frequency* of errors multiplied by the magnitude of their cost. Thus, even if the *cost* of errors remains the same, they might become more tolerable if the prize is more likely to be won. Note that the frequency of errors can depend on the environment (e.g., the distribution of the problem in space), one's own behavior (e.g., encountering the problem a lot or a little), or the behavior of others (e.g., how often they compete, help, or interfere with the problem at hand). This suggests a range of predictions for where and when we should see a bias, how strong it is, and which direction it is in.

association between the cause and effect of events. For example, because the costs of predation exceed the costs of hiding, evolution should favor mechanisms that overestimate the probability that any noise signals a predator, even when there is little or even no real link [53]. The logic and model are strikingly similar to EMT: 'natural selection can favor strategies that lead to frequent errors in assessment as long as the occasional correct response carries a large fitness benefit' [53]. Foster and Kokko made efforts to model behavioral outcomes without making any assumptions about psychological beliefs. This was largely in recognition of the difficulty of modeling psychological representations (who knows what animals really think?). However, this approach also serves to expand the model's generality; without any preconditions about cognition, the logic applies to any organism – whether bacterium, animal, or human. The effect was powerful enough for them to conclude that superstitious behaviors are 'an inevitable feature of adaptive behavior in all organisms, including ourselves' [53].

Finally, a recent model of the evolution of overconfidence [33] that contained no EMT assumptions arrived at remarkably similar conditions for the evolution of 'optimism' developed within EMT [15,28]. Overconfidence was favored as long as there was uncertainty in assessments and the relative benefits of competing for some resource were sufficiently large compared with the costs of fighting. Again, the underlying logic explaining the success of the bias is one of managing alternative possible errors to maximize fitness over time. The authors show mathematically that the evolutionary logic underlying their model yields behavior that is equivalent to a rational actor with full information who incorporates a specific 'risk premium' (e.g., see [4]) into all decisions. The chief advantages of the evolutionary model are that it: (i) shows that biases allow actors to make equally effective decisions with much less information; and (ii) explains the origin of otherwise arbitrary risk preferences by rooting them in specific environmental contexts defined by the relative impact of costly errors.

Differences and commonalities

What is striking about all of these studies (and others in Table 2), few of which cited each other, is that they identify biases as effective decision-making strategies because of the asymmetric costs of false-positive and false-negative errors made under uncertainty. This is a phenomenon that has long been recognized in the biological literature as risk sensitivity, as well as in economics as risk management [19,20,46,47,54]. The logic applies to any decision made under uncertainty with asymmetric costs – whether by animals, humans, organizations, or machines. The fact that the biases under study were very different, and in very different domains, makes these works all the more significant; the conditions for error management appear to generalize across domains and types of actors (summarized in Table 3 and Box 1).

The broad scope of error management

A framework for psychological biases

Haselton and Nettle argued that EMT accounts for various psychological biases in three broad domains [15]: (i) protective effects in perception, attention, and learning (e.g., the sound of approaching objects, bodily protection from harm, dangerous animals, dangerous people, food aversions); (ii) biases in interpersonal perception (e.g., the illusion of animacy, the sinister-attribution error, the social-exchange heuristic, sex differences in interpreting courtship signals); and (iii) self-related biases (e.g., positive illusions, the illusion of control).

Although this is an impressive list, EMT might be considerably more general. *All* psychological biases entail potential costs because they amount to false beliefs about the world (note here that we say 'bias' as it is usually used in psychology, meaning any distortion of reality). In many cases, the costs of consequent decision-making errors might be small. But any cost – however small – introduces the possibility that there is a tradeoff between the costs of false-positive and false-negative errors. If so, a bias can serve to minimize these costs. For any given decision domain, the odds that the costs of alternative errors have been identical over evolutionary time is essentially zero. Hence, error-management biases should be pervasive and the norm in the animal world, not the exception. As noted, only a small asymmetry is needed for the evolution of a bias

Table 3. Common elements in explicit mathematical models of error management

Model	Basic conditions for EMT		Costs of		Conditions for evolution of bias
	Uncertainty	Differential costs	False positive	False negative	
Green and Swets (1966); Swets (1992) [61,72]	Yes (various examples)	Yes (various examples)	Generally low	Generally high	N/A
Bouskila and Blumstein (1992) [48]	Yes (predation always possible)	Yes	Poor foraging	Predation	Large errors in estimating the risk of predation
Nesse (2005) [24]	Yes (environmental risks unpredictable)	Yes	Small metabolic costs	Illness	N/A
Haselton and Nettle (2006) [15]	Yes (social interactions hard to predict)	Yes (various examples)	Generally low	Generally high	$b/c > 1$ (where b is benefit, c is cost)
Foster and Kokko (2009) [53]	Yes (cause and effect only loosely correlated)	Yes	Wasted time	Predation	Moderate (not weak or strong) associations between events
McKay and Efferson (2010) [32]	Yes (decisions inherently noisy)	Yes (various examples; show these need not be large)	Generally low	Generally high	Small cost differentials; constraints on Bayesian updating
Johnson and Fowler (2011) [33]	Yes (relative capabilities vary)	Yes	Missed opportunity	Defeat	$b/c > 3/2$ (binary errors) $b/c > 0.7$ (continuous errors)

[32,33]. Uncertainty is also likely to be the norm rather than the exception. Few decisions, if any, are made in the absence of uncertainty. Uncertainty arises for a host of reasons, including the difficulty of predicting outcomes and how different variables might interact.

We suggest, therefore, that most if not all psychological biases can be recast in EMT logic, including major well-established biases such as ingroup–outgroup bias, threat sensitivity, and prospect theory (risk acceptance when facing losses and risk avoidance when facing gains) [4,11,14,23,55]. Many such biases were identified and examined without any explicit EMT logic, but they fit within this broader framework. Humans are far from perfect Bayesian decision makers, but natural selection has generated alternative decision-making heuristics that (approximating the Bayesian ideal) maximized fitness over evolutionary time in the presence of asymmetric costs of errors and the absence of certainty [34,35]. This diversity of applications opens up a range of complications and extensions for new research on EMT (Box 2).

A framework for phenomena beyond psychology

A primary benefit of the error management perspective is that it helps us to organize our understanding of human *psychological* biases, as outlined above. However, its underlying logic extends to any *system* that experiences differential costs of errors and uncertainty. Such systems occur at many levels in animals and humans – whether

Box 2. Complications and extensions of Error Management Theory (EMT)

Perceptions

The applications of EMT are expanded further when we consider how people *perceive* error management dilemmas. This poses some interesting problems. First, people can *misperceive the magnitude of costs* (and/or benefits). At the extreme, people might think false positives are more costly than false negatives (or vice versa), when the reverse is true. This could lead to especially damaging behavior. Second, people can *misperceive the probabilities of outcomes*. Humans are sensitive to the way in which probabilities are presented, which can radically affect their judgments and decisions [10,73]. Of particular importance, people tend to underweight the probability of events, but this tendency is reversed for very low-probability events, which are overweighted [23].

Third, there can be a *mismatch* between (real or perceived) cost asymmetries in the modern environment and those characterizing decisions during human evolutionary history. If our evolved cognitive biases emerge reliably in diverse developmental environments, they can lead us into damaging behavior somewhat irrespective of the contemporary consequences.

Multiple agents

Most of the EMT literature considers individuals making decisions in isolation. However, the problem might be magnified among two (or more) individuals both facing an error management problem. This can be illustrated with the classic human mate-search bias. Evolutionary theory suggests that males should overestimate female interest and that females should do the opposite, because of differing costs and benefits of reproductive strategies between the sexes [26,74]. However, if males are overconfident as a result, females should compensate to discount this overconfidence, becoming coyer still. Males in turn must increase their overconfidence further. Multiplayer settings might thus lead to an arms race in which biases become exaggerated (cf. [75]).

Interactions

EMT describes a solution to a very simple decision problem (do this or that within a given setting). However, in reality decisions can be complicated in several ways. For example: decision-making biases might have evolved to solve a problem in one domain but then be activated in other domains (such as ingroup–outgroup biases that evolved in small-scale societies undermining larger-scale cooperation in modern societies); multiple agents might be attempting to solve a common problem with different actual or perceived cost asymmetries (such as different states experiencing very different costs and benefits from regulations on carbon emission); and one bias might co-occur with contradictory biases acting at the same time (such as overoptimism conflicting with many forms of negative bias). Given the number of psychological biases and the complexity of social interactions, how errors are managed in the real world presents significant challenges.

genes, biochemical processes, individuals, or groups. They also occur in ecosystems, machines (e.g., smoke alarms), commercial competition, the stock market, politics, and medicine [7,24,25,56]. All such systems face the problem of managing different errors under uncertainty, which can lead to the emergence of adaptive biases – whether consciously or not – that minimize the costs of errors over time. All that is different is the unit and the mechanism of selection. So, some applications of error management reflect the basic engineering problem of error management, where humans can consciously or subconsciously arrive at a bias that results in effective decisions. Other examples reflect evolved biases that have exploited the engineering principle of error management over evolutionary time, producing biases optimized for the evolutionary past – and which therefore may not work well today.

Accounting for disciplinary differences

How have so many researchers developed essentially the same logic independently of each other? We suggest that one important reason is that different disciplines tend to focus on different aspects of Tinbergen's famous four questions about the causes of behavior: function, mechanism, phylogeny, and development [57,58]. For example, some authors have focused on function (what problem do biases solve?), whereas others have focused on mechanism (what cognitive mechanisms give rise to biases?). Perhaps unsurprisingly, it is ethologists and evolutionary biologists that have focused on function (and thus biases evident in *behavior*) (e.g., [24,48,53,59]); psychologists have focused on proximate mechanisms (and thus biases evident in *cognition*) (e.g., [31,32]) and economists have focused on a form of development (and thus biases as a result of *learning* or Bayesian updating) (e.g., [32]). Evolutionary psychologists lie somewhere in between, focusing on functional outcomes but also being necessarily interested in cognitive mechanisms and their interaction with information from the environment [15]. It might therefore be no surprise that EMT developed in the discipline of evolutionary psychology as an effort to reconcile both cognitive and behavioral patterns.

Concluding remarks

Toward a common framework

Error management offers a unifying framework for understanding decisions made under uncertainty, wherever there are asymmetric costs of false-positive and false-negative errors. Because most decisions are made with some uncertainty, and because most outcomes have different costs, error management offers considerable utility in identifying, explaining, and even recommending biases across a range of disciplines and applications. We have aimed to draw attention to the diverse but convergent recognition of the phenomenon of error management, common sources of variation in the effective management of errors, and an evolutionary explanation for cognitive biases as heuristics to exploit error management.

An evolutionary explanation for errors

The central message of EMT – that is, the application of error management principles to understand human judgment and decision making – suggests that occasional 'mistakes' are to be expected, and where they occur, they can betray adaptive, not maladaptive, behavior. An evolutionary perspective also offers ways to identify and predict cases of 'mismatch', where error management biases that evolved to deal with environments of our evolutionary past are likely to cause damaging behavior in the modern environment. The probabilities and costs of decision-making errors in any given domain are likely to be very different between the past and present.

In essence, we have a converging theory of the evolution of error. Developing a framework to study this counterintuitive logic can increase our ability to understand, and improve, the way we balance personal, social, economic, and political risks. Error management might therefore help us to avoid the costly mistakes that have plagued human history, while also taking the risks that have driven remarkable feats of human endeavor.

Acknowledgments

The authors thank Paul Craze, Lee Cronk, Tom Dickins, Charles Efferson, Melissa Fales, Ernst Fehr, Kevin Foster, Willem Frankenhuis, Kelly Gildersleeve, Jennifer Hahn-Holbrook, Rob Kurzban, Christina Larson, Aaron Lukaszewski, James Marshall, Ryan McKay, Daniel Nettle, Peter Nonacs, Daniel Rubenstein, Shimon Saphire-Bernstein, Adrienne Tecza, Pete Trimmer, Robert Trivers, and Bill von Hippel for discussion and comments contributing to this manuscript.

References

1 Wagner, A. (2003) Risk management in biological evolution. *J. Theor. Biol.* 225, 45–57
2 Evans, D. (2012) *Risk Intelligence: How to Live with Uncertainty*. Free Press
3 Alden-Smith, E. (1988) Risk and uncertainty in the "original affluent society": evolutionary ecology of resource sharing and land tenure. In *Hunters and Gatherers (History, Evolution and Social Change, Vol. I)* (Ingold, T. *et al.*, eds), pp. 222–252, Berg
4 McDermott, R. *et al.* (2008) On the evolutionary origin of prospect theory preferences. *J. Polit.* 70, 335–350
5 Rubenstein, D.I. (1982) Risk, uncertainty and evolutionary strategies. In *Sociobiology* (Bertram, B.C.R. *et al.*, eds), pp. 91–111, Cambridge University Press
6 Vermeij, G.J. (2004) *Nature: An Economic History*. Princeton University Press
7 Swets, J.A. *et al.* (2000) Psychological science can improve diagnostic decisions. *Psychol. Sci. Public Interest* 1, 1–26
8 Gilovich, T. *et al.*, eds (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press
9 Fiske, S.T. and Taylor, S.E. (2007) *Social Cognition: From Brains to Culture*. McGraw-Hill
10 Kahneman, D. (2011) *Thinking, Fast and Slow*. Allen Lane
11 Kahneman, D. *et al.* (1982) *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press
12 Akerlof, G.A. and Shiller, R.J. (2009) *Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*. Princeton University Press
13 Sears, D.O. *et al.* (2003) *Oxford Handbook of Political Psychology*. Oxford University Press
14 Fiske, S.T. (2002) What we know about bias and intergroup conflict, problem of the century. *Curr. Dir. Psychol. Sci.* 11, 123–128
15 Haselton, M.G. and Nettle, D. (2006) The paranoid optimist: an integrative evolutionary model of cognitive biases. *Pers. Soc. Psychol. Rev.* 10, 47–66
16 Buss, D.M. (2005) *The Handbook of Evolutionary Psychology*. Wiley
17 Barkow, J.H. *et al.*, eds (1992) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford University Press
18 Dunbar, R. and Barrett, L., eds (2009) *Oxford Handbook of Evolutionary Psychology*, Oxford University Press
19 Davies, N.B. *et al.* (2012) *An Introduction to Behavioural Ecology*. Wiley Blackwell

20 Bateson, M. and Kacelnik, A. (1998) Risk-sensitive foraging: decision-making in variable environments. In *Cognitive Ecology* (Dukas, R., ed.), pp. 297–341, University of Chicago Press
21 McNamara, J.M. *et al.* (2011) Environmental variability can select for optimism or pessimism. *Ecol. Lett.* 14, 58–62
22 Trimmer, P.C. *et al.* (2011) Decision-making under uncertainty: biases and Bayesians. *Anim. Cogn.* 14, 465–476
23 Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decisions under risk. *Econometrica* 47, 263–291
24 Nesse, R.M. (2005) Natural selection and the regulation of defenses: a signal detection analysis of the smoke detector problem. *Evol. Hum. Behav.* 26, 88–105
25 Nesse, R.M. (2001) Natural selection and the regulation of defensive responses. *Ann. N. Y. Acad. Sci.* 935, 75–85
26 Haselton, M.G. and Buss, D.M. (2000) Error management theory: a new perspecitive on biases in cross-sex mind reading. *J. Pers. Soc. Psychol.* 78, 81–91
27 Haselton, M.G. and Galperin, A. (2013) Error management in relationships. In *Handbook of Close Relationships* (Simpson, J.A. and Campbell, L., eds), pp. 234–254, Oxford University Press
28 Nettle, D. (2004) Adaptive illusions: optimism, control and human rationality. In *Emotion, Evolution and Rationality* (Evans, D. and Cruse, P., eds), pp. 193–208, Oxford University Press
29 Yamagishi, T. *et al.* (2007) The social exchange heuristic: managing errors in social exchange. *Ration. Soc.* 19, 259–292
30 Johnson, D.D.P. (2009) The error of God: error management theory, religion, and the evolution of cooperation. In *Games, Groups, and the Global Good* (Levin, S.A., ed.), pp. 169–180, Springer
31 McKay, R.T. and Dennett, D.C. (2009) The evolution of misbelief. *Behav. Brain Sci.* 32, 493–561
32 McKay, R.T. and Efferson, C. (2010) The subtleties of error management. *Evol. Hum. Behav.* 31, 309–319
33 Johnson, D.D.P. and Fowler, J.H. (2011) The evolution of overconfidence. *Nature* 477, 317–320
34 Gigerenzer, G. and Todd, P.M. (2000) *Simple Heuristics That Make Us Smart*. Oxford University Press
35 Cosmides, L. and Tooby, J. (1994) Better than rational: evolutionary psychology and the invisible hand. *Am. Econ. Rev.* 84, 327–332
36 Haselton, M.G. and Buss, D.M. (2009) Error management theory and the evolution of misbeliefs. *Behav. Brain Sci.* 32, 522–523
37 Sharot, T. *et al.* (2011) How unrealistic optimism is maintained in the face of reality. *Nat. Neurosci.* 14, 1475–1479
38 Sharot, T. (2011) The optimism bias. *Curr. Biol.* 21, R941–R945
39 Marshall, J.A.R. *et al.* (2013) On evolutionary explanations of cognitive biases. *Trends Ecol. Evol.* 28, 469–473, this issue
40 Kahneman, D. and Tversky, A. (1982) *Subjective probability: a judgment of representativeness In Judgment under Uncertainty: Heuristics and Biases*. pp. 32–47, Cambridge University Press
41 Haselton, M.G. *et al.* (2009) Adaptive rationality: an evolutionary perspective on cognitive bias. *Soc. Cogn.* 27, 733–763
42 Finkelman, F.D. (2012) The price we pay. *Nature* 484, 459
43 Martincorena, I. *et al.* (2012) Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature* 485, 95–98
44 Martin, T.L. and Huey, R.B. (2008) Why "suboptimal" is optimal: Jensen's inequality and ectotherm thermal preferences. *Am. Nat.* 171, E102–E118
45 Dawkins, R. and Krebs, J.R. (1979) Arms races between and within species. *Proc. R. Soc. Lond. B: Biol. Sci.* 205, 489–511
46 Stephens, D.W. and Krebs, J.R. (1986) *Foraging Theory*. Princeton University Press
47 Stephens, D.W. *et al.*, eds (2007) *Foraging: Behavior and Ecology*, University of Chicago Press
48 Bouskila, A. and Blumstein, D.T. (1992) Rules of thumb for predation hazard assessment: predictions from a dynamic model. *Am. Nat.* 139, 161–176
49 Abrams, P.A. (1994) Should prey overestimate the risk of predation? *Am. Nat.* 144, 317–328
50 Wiley, R.H. (1994) Errors, exaggeration, and deception in animal communication. In *Behavioral Mechanisms in Ecology* (Real, L., ed.), pp. 157–189, University of Chicago Press
51 Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* 103, 193–210
52 Sharot, T. (2011) *The Optimism Bias: A Tour of The Irrationally Positive Brain*. Pantheon
53 Foster, K.R. and Kokko, H. (2009) The evolution of superstitious and superstitious-like behaviour. *Proc. R. Soc. B: Biol. Sci.* 276, 31–37
54 Bernstein, P.L. (1998) *Against the Gods: The Remarkable Story of Risk*. Wiley
55 Neuberg, S.L. *et al.* (2011) Human threat management systems: self-protection and disease avoidance. *Neurosci. Biobehav. Rev.* 35, 1042–1051
56 Tetlock, P.E. (1998) Social psychology and world politics. In *Handbook of Social Psychology* (Gilbert, D. *et al.*, eds), pp. 868–912, McGraw Hill
57 Tinbergen, N. (1963) On aims and methods in ethology. *Z. Tierpsychol.* 20, 410–433
58 Scott-Phillips, T.C. *et al.* (2011) Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspect. Psychol. Sci.* 6, 38–47
59 McNamara, J.M. and Houston, A.I. (2009) Integrating function and mechanism. *Trends Ecol. Evol.* 24, 670–675
60 Pascal, B. (2008/1669) *Pensées and Other Writings*. Oxford University Press
61 Green, D.M. and Swets, J.A. (1966) *Signal Detection Theory and Psychophysics*. Wiley
62 Egan, J.P. (1975) *Signal Detection Theory and ROC Analysis*. Academic Press
63 Higgins, E.T. (1997) Beyond pleasure and pain. *Am. Psychol.* 52, 1280–1300
64 Pacala, S.W. *et al.* (2003) False alarm over environmental false alarms. *Science* 301, 1187–1188
65 McNamara, J.M. and Houston, A.I. (1985) The choice of two prey types that minimizes the probabilty of starvation. *Behav. Ecol. Sociobiol.* 17, 135–141
66 Loewenstein, G. *et al.*, eds (2003) *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, Russell Sage Foundation
67 Pennebaker, J.W. *et al.* (1979) Don't the girls get prettier at closing time: a country and western application to psychology. *Pers. Soc. Psychol. Bull.* 5, 122–125
68 Koops, M.A. and Abrahams, M.V. (1998) Life history and the fitness consequences of imperfect information. *Evol. Ecol.* 12, 601–613
69 Caraco, T. *et al.* (1990) Risk-sensitivity: ambient temperature affects foraging choice. *Anim. Behav.* 39, 338–345
70 Houston, A.I. and McNamara, J.M. (1988) Fighting for food: a dynamic version of the Hawk-Dove game. *Evol. Ecol.* 2, 51–64
71 Enquist, M. and Leimar, O. (1990) The evolution of fatal fighting. *Anim. Behav.* 39, 1–9
72 Swets, J.A. (1992) The science of choosing the right decision threshold in high-stakes diagnostics. *Am. Psychol.* 47, 522–532
73 Cosmides, L. and Tooby, J. (1996) Are humans good intuitive statisticians after all? Rethinking some conclusions of the literature on judgment under uncertainty. *Cognition* 58, 1–73
74 Perilloux, C. *et al.* (2012) The misperception of sexual interest. *Psychol. Sci.* 23, 146–151
75 Trivers, R.L. (2011) *Deceit and Self-Deception: Fooling Yourself the Better to Fool Others*. Allen Lane