ARTICLE IN PRESS
Consciousness and Cognition xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Can evolution get us off the hook? Evaluating the ecological defence of human rationality [☆]

Maarten Boudry [a,*], Michael Vlerick [b], Ryan McKay [c]

[a] Ghent University, Department of Philosophy & Moral Sciences, St.-Pietersnieuwstraat 49, Room 204, 9000 Ghent, Belgium
[b] University of Johannesburg, Department of Philosophy, PO Box 524, Auckland Park 2006, South Africa
[c] Royal Holloway, University of London, ARC Centre of Excellence in Cognition and its Disorders, Department of Psychology, Egham, Surrey TW20 0EX, United Kingdom

ARTICLE INFO

Article history:
Received 13 June 2014
Revised 26 August 2014
Accepted 27 August 2014
Available online xxxx

Keywords:
Adaptation
Cognitive bias
Ecological rationality
Error management
Evolutionary psychology
Fallacies
Fast & frugal heuristics
Irrationality
Locus shifts

ABSTRACT

This paper discusses the ecological case for *epistemic innocence*: does biased cognition have evolutionary benefits, and if so, does that exculpate human reasoners from irrationality? Proponents of 'ecological rationality' have challenged the bleak view of human reasoning emerging from research on biases and fallacies. If we approach the human mind as an adaptive toolbox, tailored to the structure of the environment, many alleged biases and fallacies turn out to be artefacts of narrow norms and artificial set-ups. However, we argue that putative demonstrations of ecological rationality involve subtle locus shifts in attributions of rationality, conflating the adaptive rationale of heuristics with our own epistemic credentials. By contrast, other cases also involve an ecological reframing of human reason, but do not involve such problematic locus shifts. We discuss the difference between these cases, bringing clarity to the rationality debate.

1. Introduction

Like any other biological organ, the human brain is a product of evolution by natural selection: The brain secretes thought as the liver secretes bile, wrote the 18th century French physiologist Pierre Cabanis in his *Des Rapports du Physique et du Morale de l'Homme*. But does the brain's biological provenance suggest it must be successful at producing true beliefs and making rational judgments? Several generations of psychologists have suggested otherwise, documenting the myriad flaws and foibles of human reasoning (Gilovich, Griffin, & Kahneman, 2002; Kahneman, 2011; Kahneman, Slovic, & Tversky, 1982). Especially in popular summaries of this research, humans come off rather badly – prone to all sorts of biases and fallacies, woefully inadequate at dealing with probability and uncertainty, and inclined to persist in making errors of social judgment, even after these have been clearly spelled out (Ariely, 2009; Piattelli-Palmarini, 1996; Shermer, 2011; Singer & Benassi, 1981; Sutherland, 2007). The psychologist John Kihlstrom (2004) has called this the 'People are Stupid' school of psychology (PASSP).

☆ This article is part of a special issue of this journal on Imperfect Cognitions.
* Corresponding author.
E-mail address: maartenboudry@gmail.com (M. Boudry).

http://dx.doi.org/10.1016/j.concog.2014.08.025
1053-8100/© 2014 Published by Elsevier Inc.

Recently, however, a new wave of research, inspired by evolutionary ideas, has posed a challenge to this bleak picture of human reason. This school of thought heralds an ecological conception of rationality, aligning itself with research in evolutionary psychology. If we approach the human mind as a collection of cognitive heuristics, tailored to the structure of the ecological environment, many alleged biases and fallacies arguably emerge as artefacts of narrow norms and artificial set-ups.

After introducing this 'ecological rationality' research program, we identify a 'locus shift' in some attributions of rationality. In these cases, individual reasoners are being praised as rational simply because the heuristics employed in their reasoning show adaptive design (i.e. are "rational" from an evolutionary perspective). In other words, the adaptive rationale of cognition (evolutionary locus) is conflated with our own epistemic credentials (personal locus). We then evaluate whether the normative categories of (ir)rationality can still be applied at the evolutionary level of analysis, and conclude that this facet of the 'ecological defence' is confusing: (a) human reasoners do not deserve the "credit" for the adaptive designs they have been equipped with, and (b) evolution cannot exculpate human irrationality. Finally, we discuss cases in which the program of ecological rationality has succeeded in rehabilitating human reason, showing that earlier charges of irrationality had been premature. These cases also involve a form of ecological reframing, but they do not involve adaptive locus shifts.

The aim of this paper is to provide some much needed clarification to the debate about rationality. In particular, by clarifying which 'defences' of putative irrationality are legitimate and which illegitimate, we hope to provide a robust conceptual framework within which to situate and evaluate future empirical evidence. The debate about rationality is not merely of academic interest: irrationality is a topic of critical personal and social import. Biased beliefs about the self and the future may promote individually harmful behaviours like smoking, unsafe sex and overspending, as well as potentially precipitating global catastrophes such as sectarian violence, world wars, exploding financial bubbles and environmental disasters (Johnson & Fowler, 2011; Sharot, 2011). Given these wide-ranging outcomes, getting clear about the nature and extent of human irrationality is a critical philosophical and psychological project.

2. Ecological rationality

2.1. Recasting rationality in the environment

In the heuristics and biases program instigated by Daniel Kahneman and Amos Tversky in the 80s (Kahneman, 2011; Kahneman et al., 1982), heuristics are viewed as mental short-cuts leading to imperfect and often outright irrational inferences. Though Kahneman and Tversky pointed out that even fallible heuristics often lead to successful inferences, over time this sense of balance was lost, and a negative slant began to dominate the research program they founded (Krueger & Funder, 2004).[1] The program of ecological rationality, by contrast (Fawcett et al., 2014; Gigerenzer, Hertwig, & Pachur, 2011; Gigerenzer & Todd, 1999; Hertwig & Hoffrage, 2013; Todd & Gigerenzer, 2012), aims to effect a gestalt switch, urging us to rethink the norms of what counts as rational.

The canons of 'classical' rationality, to which Kahneman & Tversky's subjects were held accountable, consist of a general, formal, content-free framework for valid reasoning (e.g. modus ponens). Advocates of 'ecological rationality', by contrast, uphold a radically different view of heuristic reasoning. Heuristics, they argue, provide us with an 'adaptive toolbox' (Gigerenzer, 2008; Gigerenzer & Todd, 1999), each suited to a particular set of challenges endemic to a particular environment. In contrast with unbounded models of classical rationality, which typically assume unlimited resources both with regard to information gathering and computational processing, ecological rationality is 'fast and frugal'. According to Gigerenzer and Todd (1999, p. 13), a heuristic is ecologically rational 'to the degree that it is adapted to the structure of an environment'. Heuristics are quick and computationally cheap, requiring few and simple computational steps, and operating on a limited input domain. For example, when confronted with two objects, only one of which is recognized, people infer that the familiar one will be more important or have a greater value. Although this so-called 'recognition heuristic' – on which more later – leads us astray in artificial set-ups, it appears to be a surprisingly accurate guide to real world problems (Goldstein & Gigerenzer, 2002).

Proponents of ecological rationality claim that many apparent demonstrations of irrationality in the heuristics and biases program are artefacts of inappropriate standards and narrow norms, such as coherence criteria and the axioms of probability theory. Indeed, many experiments are expressly designed to mislead or distract participants, placing them in situations where their usually successful strategies and heuristics lead them astray. If one canvasses the same phenomena in a broader framework, however, taking into account the structure of real environments, computational limitations and various trade-offs, the human mind emerges as more rational than many psychologists suppose.

2.2. Two strands

We detect two major strands in the re-appreciation of (apparent) human irrationality as ecologically rational. On the one hand, as the ecological rationalist points out, traditional cognitive psychologists have too often delighted in tripping up their subjects with artificial set-ups that truncate the complexity of real life and human intelligence (first strand). On the other

[1] In his latest work, Kahneman is at pains to disavow this pessimistic turn: "I often cringe when my work with Amos is credited with demonstrating that human choices are irrational, when in fact our research only showed that humans are not well described by the rational-agent model" (Kahneman, 2011, p. 411).

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

hand, ecological rationalists have argued that heuristics are *not designed* to perform well in artificial (non-ecologically valid) environments. If putatively 'irrational' subjects were tested in an ecologically relevant context, the kind to which their evolved heuristics are attuned, they would behave rationally (second strand).

In both strands, ecological rationality focuses our attention on the match between our reasoning and the real world. The sensible argument behind this is that we should evaluate heuristics in their 'proper' context. You should not use a screwdriver as a crowbar and be surprised that it breaks.[2] Whereas the first strand focuses on artificial set-ups and uncharitable experimenters, the second strand invokes evolutionary rationales. In the latter case, advocates of ecological rationality point out that heuristics are crafted by evolution to deal with particular adaptive problems in an ancestral environment. Therefore, they have a "proper domain" of application (Sperber, 1996) – what Millikan (1984) termed "Normal conditions" – outside of which we should not expect them to perform well. Blaming them for "malfunction" in unnatural contexts, without proper appreciation of their design features, is like blaming the screwdriver. Nevertheless, as we will argue, while the former approach is (often) warranted, the latter involves problematic *locus shifts*.

2.3. Evolutionary psychology

This second evolutionary strand runs through much of the literature on ecological rationality, (Gigerenzer & Selten, 2002). And indeed, the program connects with work in the field of evolutionary psychology on the adaptive value of cognitive bias (Aktipis & Kurzban, 2004; Cosmides & Tooby, 1994; Haselton & Nettle, 2006; Haselton, Nettle, & Andrews, 2005; Mercier & Sperber, 2011; Nesse, 2001), the main thrust of which is also to show the ecological validity of human reason. As Haselton & Nettle write:

> Some cognitive illusions disappear or greatly attenuate when the task is presented in an ecologically valid format (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Ecological validity, a long-standing but undertheorized term in psychology, may in effect be equated to the task format approximating some task that humans have performed recurrently over evolutionary time.
>
> [Haselton & Nettle, 2006, p. 63]

Apparent biases and illusions are recast in an evolutionary light, as serving some adaptive rationale. From an adaptationist point of view, human reason is not the botched and bias-riddled device many psychologists take it to be, but emerges as an extremely effective, highly specialized set of adaptive tools for dealing with recurring problems in the Environment of Evolutionary Adaptedness (EEA).

In an overview of the research program of ecological rationality, Gerd Gigerenzer wrote that 'what appears to be a fallacy can often also be seen as adaptive behaviour, if one is willing to rethink the norms' (Gigerenzer, 2008, p. 13). In a similar vein, Haselton and Nettle (2006, p. 59) have argued that 'bias in cognition is no longer a shortcoming in rational behaviour, but an adaptation of behaviour to a complex, uncertain world'. Elsewhere, Haselton and colleagues wrote 'an adaptationist perspective suggests that the mind is remarkably well designed for important problems of survival and reproduction, and not fundamentally irrational.' These assessments seem to provide solace for the pessimistic view of human reason voiced by many psychologists. But can evolution really get us off the hook?

3. Assessment of the locus shift in ecological rationality

3.1. Minimal conditions for rationality

In order to assess what exactly ecological/adaptive rationality amounts to, we need a serviceable conception of rationality. While attempting to formulate a full blown definition seems doomed to fail, given the loose and varied use of the term, some constraints on the meaning of 'rationality' will guide us through this treacherous terrain. In this paper, we will be mainly concerned with epistemic rationality (belief formation that is accurate and truth-tracking) and not so much with instrumental rationality (actions that maximize the probability of success, according to some standard). According to De Sousa (2009, p. 289), rationality presupposes intentionality. An inanimate object or mere mechanism cannot be rational. It would be absurd to claim that a rock falling to the ground is rationally obeying the laws of gravity, much as it would be strange to call a calculator rational because it yields correct answers. Likewise, *ir*rationality is a concept that can be meaningfully applied only to intentional creatures – entities *capable* of full-blown rationality (de Sousa, 2007). Irrationality is a term of epistemic disapprobation, used when someone fails to live up to some expected normative standard. As *ought* implies *can*, it makes no sense to blame a mindless process:

> It makes sense to criticize a person, but not a cell, for having made a mistake in computation, or with having failed to foresee what should have been foreseen, or with having acted on reasons that fell short of the best set of reasons.
>
> [de Sousa, 2007, p. 11]

[2] There is a range of possibilities as to the selection of heuristics. They can be consciously crafted and selected to tackle certain problems (through individual learning) or, they can be crafted by evolution though applied consciously, or learned (reinforced) through habit and applied unconsciously, etc. (see Marewski & Schooler, 2011; Rieskamp & Otto, 2006).

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

In an extended, *metaphorical* sense, however, standards of rationality can be fruitfully applied to the domain of functionality, including biological adaptations and artefacts, even when no intentionality is involved. Although natural selection is a mindless process, it can produce an *appearance* of intentionality by crafting design solutions to recurring adaptive problems (Williams, 1966/1996). Evolution by natural selection gives rise to what Dennett coined 'free-floating rationales' (Dennett, 1983, p. 351). In this extended sense of rationality, the relevant measure of success is equated with reproductive success or inclusive fitness (Over, 2000), and standard frameworks of rationality such as game theory (Maynard Smith, 1974) or expected utility theory (see Section 3.2) can be applied (Kacelnik, 2006).

Because of the absence of intentionality, however, not all dimensions of full-blown rationality can be transferred to evolution. Whereas individual reasoners are capable of foresight, evolution cannot look into the future. For example, selection can easily get stuck on local peaks in an adaptive landscape, unable to reach nearby higher summits. If evolution had the capacity of foresight, it would backtrack and make a detour, as a human reasoner would. Evolution, therefore, provides a mere 'simulacrum of rationality' (de Sousa, 2009, p. 298).

3.2. Locus shifts

Given that rationality (in the literal sense) can only be attributed to person-level decision making (see above), we question whether the ecological reframing succeeds in dispelling the charge of irrationality. In order to get clear about the notion of 'locus shifts', we first discuss an example from evolutionary psychology in which a similar calculus of rationality may be implemented either at the individual level or on an evolutionary scale: error management. This will allow us to pull apart the two loci of rationality, while holding the measure of success (e.g. protection from bodily harm) constant.

Expected utility theory is one of the hallmarks of classical rationality: it describes how to maximize utility under uncertainty, as a function of the probability of different outcomes and their associated costs and benefits. Normally, the model is used to describe individual choices, for example betting preferences. In recent years, however, the logic of expected utility theory has been extended to evolutionary adaptations, under the guise of error management theory (Haselton & Buss, 2000; Haselton & Nettle, 2006; Nesse, 2001).[3] In a range of different situations, organisms need to process ambiguous stimuli that may or may not signal adaptively relevant situations, such as the presence of a predator or pathogen. If the costs associated with the respective errors, measured in terms of (inclusive) reproductive fitness, are asymmetric, then expected utility theory can be applied to predict optimal behaviour. If the costs associated with one type of error are larger than those associated with the opposite kind, then a strategy biased in favour of making the less costly error may pay off in the long run, even if that increases the absolute numbers of errors (Arkes, 1991). As Haselton et al. put it: 'it is better to make more errors overall as long as they are of the relatively cheap kind' (Haselton et al., 2005, p. 731). Error management is displayed in a range of biological adaptations that are clearly beyond our voluntary control, such as inflammation, disgust, fever and coughing. In all of these cases, the locus of rationality is clearly evolution by natural selection, not the individual subject.

Things get more interesting when considering applications of error management theory to human cognition. In such cases, the locus of rationality may reside either at the individual level or at the evolutionary level. In most discussions of error management, human behavioural biases are interpreted in terms of biased belief, with evolution as the rational bookkeeper. For example, people buy into various superstitions because they evolved pattern detection modules that err on the side of caution. In order not to miss out on any important causal patterns in the world, evolution has 'decided' to run the risk of occasional false positives, which are mostly relatively harmless.

In a commentary on error management theory, however, McKay and Efferson (2010) argued that human behavioural bias does not automatically entail biased belief. In principle, the same behavioural bias may be arrived at through accurate belief formation combined with judicious action policies, based on expected utility theory. If I react to a rustle in the leaves as if there is a predator lurking in the bushes, I do not need a strong *belief* that a predator is there. Likewise, when approaching a girl that I like, I do not need to be certain that she fancies me. I may decide to act upon an admittedly remote possibility, in light of the high opportunity costs associated with false negatives (ending up as lunch, or missing a romantic – read reproductive – opportunity). As McKay & Efferson noted, 'there are infinitely many ways to accomplish the required behavioural change ... inferences about cognition can be radically underdetermined when one observes an interesting behavioural bias.' (see also Marshall, Trimmer, Houston, & McNamara, 2013; McKay & Efferson, 2010, p. 312).

The logic underlying the behavioural bias is the same, but our normative appraisal of human reasoners depends on where the *locus* of rationality resides. If the behavioural bias is the outcome of a prudential action policy, without the involvement of biased belief, we should not accuse reasoners of irrationality. For example, people fasten their seatbelts before starting the car not because they strongly believe they will crash into other cars, but just because they would rather be safe than sorry: the cost of a false negative (being flung through the windscreen) is orders of magnitude larger than that of a false positive (the energy expended in fastening one's seatbelt and the mild discomfort of being strapped in it), large enough to offset the small probability of a serious car accident. This is a case of full-blown (intentional) rationality, situated at the level of individual decision making. An alien scientist interpreting this prudential human behaviour as a symptom of irrational paranoia would be very uncharitable indeed. On the other hand, if a behavioural bias stems from some attendant belief distortion, the

[3] Here we ignore the point that error management theory is based on optimization, which would be a problematic form of unbounded rationality in Gigerenzer's framework. However, in this section we are simply trying to pry apart two *loci* of rationality, and so we can ignore different conceptions of rationality for the time being.

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

charge of irrationality may be apposite. For example, if an unattractive man approaches every woman at the bar, believing that he is irresistible to the opposite sex, we may be justified in calling him irrational.

If biased behaviour translates to bias in belief formation, however, with evolution keeping the books of fitness costs and benefits, we think our assessment should be different (for a discussion, see Galperin & Haselton, 2012). From the adaptationist point of view, the ecological 'problem' of exploiting causal regularities in an organism's environment can be solved by different means: a fixed and genetically encoded action sequence, the capacity to form conditional reflexes and make associations, a belief formation system with a pre-programmed belief bias, or the genuine capacity for causal thought and expected utility reasoning (Dennett, 1996; Lorenz, 1941/2009). Only in the latter case is rationality displayed at the locus of the individual decision-maker. For example, many people engage in superstitious rituals because they strongly believe in the causal connection. If error management theorists are right, there is indeed method in the madness of such superstition, since the fitness cost of overlooking causal links outweighs the fitness cost of detecting non-existent causal relationships. In such cases, however, the locus of 'rationality' pertains to the adaptive design of our reasoning faculties. Although it is tempting to collapse both levels of rationality, as though the adaptive rationale of our superstitious behaviour redounds to the human reasoner, the 'method' of natural selection does not obviate – indeed it requires – the 'madness' of the individual. Accordingly, we defend the following claims:

(a) We should not credit human reasoners simply because the adaptive design of their reasoning faculties conforms to some standard principles of rationality.[4] We do not congratulate people for their sophisticated immune systems or ingenious thermoregulation, nor do we credit cicadas for having discovered prime numbers,[5] or spiders for the beautiful geometry of their webs. There is neither cognitive transparency – the organism has no clue why it behaves as it does – nor intentional action. According to Dennett, If we discover that an animal is too simple-minded to harbour an adaptive rationale, we do not discard the rationale but are simply forced to 'pass the rationale from the individual to the evolving genotype' (Dennett, 1983, p. 351).

(b) Conversely, if we have good grounds for calling someone's beliefs irrational (see further), then *merely* pointing out the larger evolutionary rationale of the faculties generating those beliefs will not get him off the hook. Having biologically adaptive cognition is *not* an alibi where the charge of irrationality is concerned. If superstitious belief is adaptive, because it motivates 'better safe than sorry' behaviours, that does not make superstitious people rational. For example, while people who believe that black cats are harbingers of bad luck may be heeding to the 'wisdom' of time-honoured cognitive adaptations, that does not make their belief rational or truth-tracking.

3.3. The rationality of heuristics?

3.3.1. Statistical rationales

In the classical heuristics-and-biases framework, heuristics are perceived as imperfect shortcuts in an uncertain world, strategies that execute a trade-off between cost and accuracy. Gigerenzer and his colleagues, however, have documented several cases where this trade-off can be escaped, and 'less information and computation lead to more accurate judgments' (Gigerenzer & Sturm, 2012, p. 261). The showpiece of these 'less is more' effects is the recognition heuristic, where people take advantage of their own ignorance to make intelligent inferences. In judging which out of two cities has a larger population, for example, people use name recognition as a cue for population level. When the task consists of German cities, American subjects perform better than their German colleagues, because the latter recognize too many cities to exploit their own ignorance. A measure of ignorance turns out to be profitable in this case (Goldstein & Gigerenzer, 2002), while knowing too much is a burden. Apparently, everyday life presents us with many situations in which the recognition heuristic proves successful. It provides us with a powerful and accurate tool for making inferences about the environment.

Fast and frugal heuristics escape a trade-off between accuracy and speed because 'they make a trade-off on another dimension: that of generality versus specificity' (Gigerenzer & Todd, 1999, p. 18). In later work, Gigerenzer and Brighton (2009) have explored another dilemma (between bias and variance) for explaining when and why such 'less-is-more' effects occur. The upshot of this statistical analysis is that the simplicity of fast and frugal heuristics (few computational steps and few parameters) protects them against overfitting (i.e. mistaking noise for underlying patterns) and hence ensures robustness in the face of environmental change.

The success of the recognition heuristic derives from a probabilistic rationale. In order for the trick to work, there needs to be a correlation between the probability that an item is recognized and the value of interest. This 'recognition validity' needs to outweigh the 'knowledge validity', which is the probability of giving the correct answer when both items are recognized (Goldstein & Gigerenzer, 1999, 2002). In other words, your ignorance must be more valuable than your knowledgeability. In addition, success depends on the degree of uncertainty, the number of alternatives, and the size of the learning sample. Goldstein & Gigerenzer conclude:

[4] Nor, of course, should we criticise people for the limitations of their automatic, non-reflective cognition.

[5] Cicadas have developed a life cycle with prime-number year intervals (e.g. 17 years), because that makes it difficult for predating species to 'track' them. If they had a life cycle of 15 years, say, predators could show up every 3 or 5 years, never missing their prey. With a 17-year interval, however, predators with a 5-year life cycle catch the cicadas only once every 80 years.

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

> [M]any scholars, psychologists included, have mistrusted the power of these heuristic principles, and saw in them simple-mindedness and irrationality. This is not our view. The recognition heuristic is not only a reasonable cognitive adaptation because there are situations of limited knowledge in which there is little else one can do. It is also adaptive because there are situations . . . in which missing information results in more accurate inferences than a considerable amount of knowledge can achieve.
>
> [Gigerenzer & Todd, 1999, p. 58]

Now what should we make of this contrast between simple-mindedness and 'reasonable cognitive adaptation'? Can we attribute rationality to heuristic-wielding human thinkers when their heuristics yield accurate inferences?

3.3.2. Cognitive transparency

The answer to this question depends on the criteria of cognitive transparency and intentionality. As pointed out, attribution of (literal) rationality requires intentional deliberation on the part of the actor, not just the execution of hard-wired action patterns or mindless heuristics. (cf. 3.1 Minimal conditions for rationality). To avoid regress of justification, of course, we should not require that the agent has contemplated every step in the chain of belief-formation. Perceptual beliefs may arise automatically, without deliberation and thus without rationality, but more complex beliefs can be formed consciously on the basis of basic perceptual beliefs. On the latter level, attributions of rationality and irrationality can be meaningfully made.[6]

Without delving into empirical details, we can be confident that many people are ignorant of the statistical principles underlying the success of the recognition heuristic. This is hardly surprising if the heuristic is a 'cognitive adaptation', as Goldstein and Gigerenzer (2002) suggest. Many cognitive modules in our brain perform their proper function without us having the slightest awareness of their operation, let alone their design features and the reasons for their success. If the recognition heuristic is an evolved adaptation, activated under appropriate circumstances by a mechanism to which human cognizers have no access, it would be strange to credit people for its success. The locus of rationality, in this case, is at the level of adaptation. Evolution, rather than human cognizers, has exploited recurrent statistical correlations in the environment.

However, recent research shows that some people who use the heuristic articulate a justification that captures the ecological validity (Gigerenzer, 2007, pp. 125–126), and they will suspend the use of the heuristic if they have good counter-indications (Pachur, Todd, Gigerenzer, Schooler, & Goldstein, 2011). Even if such reasoners have no full understanding of the statistical principles involved, it seems that the criterion of transparency is satisfied, and it would be uncharitable to deny them rationality at the personal level.

In general, however, the evidence points in the other direction

> Research suggests that people hardly ever make conscious decisions about which heuristic to use but that they quickly and unconsciously tend to adapt heuristics to changing environments, provided there is feedback.
>
> [Gigerenzer, 2008, p. 38]

In another striking example from the catalogue of fast and frugal heuristics, both humans and dogs use the following rule of thumb to catch a ball in flight: keep your gaze fixed at the ball, and adjust your running speed such that the angle of the ball in your visual field remains constant (Gigerenzer & Todd, 1999). If baseball players are asked to predict where a ball will land, but are not allowed to run towards it, they perform very poorly (Babler & Dannemiller, 1993). However, when asked how they proceed to catch a ball on the fly, people are typically oblivious of using the heuristic: 'most fielders are blithely unaware of the gaze heuristic, despite its simplicity' (Gigerenzer, 2007, p. 11). This, it should be emphasised, does not mean that heuristic-wielding humans are *irrational*, but that the normative categories of (ir)rationality simply do not apply insofar as humans are working on automatic pilot.

Interestingly, on some occasions, Gigerenzer & Todd explicitly credit evolution for the design of heuristics: 'evolution would seize upon informative environmental dependencies such as this one and exploit them with specific heuristics if they would give a decision-making organism an adaptive edge' (Gigerenzer & Todd, 1999, p. 19). Such statements, however, are in tension with the main thrust of their research programme, which is to dispel accusations of 'simple-mindedness' and irrationality by pointing to the larger ecological picture. As we saw in the example of error management, adaptive design is perfectly compatible with dumbness.

In this regard, the novel insights gleaned from Gigerenzer et al.'s research program conflict with their aspiration to rehabilitate human reason. Indeed, if the rationale of the recognition heuristic or the gaze heuristic were transparent to those profiting from it, the results of Gigerenzer et al. would not have been so informative. The effectiveness of these heuristics is surprising even to the researchers themselves. For example, another deceptively simple heuristic, 'take the best', which takes the first discriminative cue between two items and ignores the other cues, has outperformed weighted statistical decision procedures such as multiple regression. This is how Gigerenzer & Goldstein relate their amazement at their own result: 'When we first obtained these results, we could not believe them.' (Gigerenzer & Todd, 1999, p. 89).

[6] There is no reason to assume that this is an all-or-nothing affair: there may be degrees of transparency and conscious deliberation.

Because people use their heuristics without cognitive transparency, they are also often insensitive to local conditions where their heuristics lead them astray.[7] For example, many advertisers exploit the recognition heuristic by investing in brand recognition rather than product quality, thereby undermining the ecological validity of the correlation between renown and product quality and fooling consumers. In the 'overnight fame' experiments by Jacoby, Kelley, Brown, and Jasechko (1989), people who had been presented with unknown names the day before were tricked into believing that those names were famous. This should not come as a surprise. Evolution exploits recurring statistical properties of the environment of evolutionary adaptiveness (EEA), but it is not responsive to any environmental contingency. For evolution it is 'rational' to implement the recognition heuristic in an organism as long as overshooting or misfiring has been sufficiently rare or inconsequential in the EEA. Error management on the evolutionary scale abstracts away from local contingencies.

3.4. Why adaptive 'rationality' does not align with personal rationality

In the section on error management, we noted that adaptively 'rational' cognition may translate into irrational belief, such as superstition. There are two additional reasons why evolutionary rationality and personal rationality may diverge. The first is the so-called problem of adaptive mismatch. If heuristics are evolved adaptations for managing cost asymmetries in the EEA, there is no guarantee that the same balance sheet still applies to modern environments. Probabilities may have changed over time, as may have the respective costs and benefits. This mismatch hypothesis is one of the main tenets of evolutionary psychology (Pinker, 1997; Tooby & Cosmides, 1994).[8] A heuristic may succeed in the EEA but fail in modern environments, or only match some limited aspects of modern environments (e.g. to predict which city will have the largest population).

A second reason concerns the issue of who – or rather what – is the ultimate beneficiary of natural selection (cui bono? see Dennett, 1995). As Dawkins (1976) has pointed out, following up on pioneering work by Williams and Hamilton (Hamilton, 1963; Williams, 1966/1996), genes can be seen as the ultimate beneficiaries of evolutionary adaptation. Individual organisms are best seen as 'vehicles' for genetic propagation (Dawkins, 1982). Take, for instance, the suicidal sting of a bee. Such behaviour is 'rational' from the point of the evolutionary processes that gave rise to such behaviour, because the bee shares 75% of its genes with its sisters. Sacrificing its life to enhance the survival chances of the hive therefore makes sense from a gene-centric perspective. But it certainly does not benefit the individual bee. Evolutionary rationality – which is situated at the level of the replicators – therefore does not always coincide with rationality at the level of the vehicle (Skyrms, 2000; Stanovich & West, 2000; Stanovich & West, 2003).

4. Genuine fallacies

Let us now have a closer look at two famous examples from the irrationality literature, where advocates of ecological rationality have performed a locus shift: the gambler's fallacy and the hot hand fallacy. The former refers to the belief of roulette players that numbers which have not turned up yet are more likely to win in the next round, because they are 'overdue'. In reality, of course, each turn of the roulette wheel is independent of the next. The wheel has no memory, so the probability distribution remains the same at every trial. More generally, the gambler's fallacy describes the belief that statistical deviations in a given direction will soon be counterbalanced by deviations in the opposite direction.

The hot hand fallacy can be seen as the mirror image of the gambler's fallacy. To commit the hot hand fallacy is to believe that, if some deviation from chance has been observed, further deviations in the same direction are more likely on subsequent trials. The name derives from basketball, where there is a widespread belief among players and fans that players score in streaks (Gilovich, 1983; Gilovich, Vallone, & Tversky, 1985). If one player has already scored well, he has a 'hot hand' and is more likely to make further points during the rest of the game. If he makes a bad start, however, there is a jinx on him and he will be less likely to perform well during the rest of the match. Statistical analyses of basketball scoring data have revealed no such clustering (but see Raab, Gula, & Gigerenzer, 2012, for evidence of streakiness in volleyball), but belief in the hot hand persists.[9]

The fact that each of these forms of inference has been bestowed its proper label in the literature, despite their being one another's mirror image, is a source of amusement to Gigerenzer and his colleagues. For them, it exemplifies the bad habit of psychologists to label any observed deviation from accepted rationality models as a 'fallacy' or 'bias', without any understanding of the underlying cognitive mechanisms (Gigerenzer & Brighton, 2009).

[7] Intelligent choice in the face of contingencies is one of the hallmarks of rationality. Perhaps people do not fully understand why the heuristics they employ work, but they may know when to implement them. In particular, they may make flexible use of the different tools in their mental toolkit, switching strategies in the face of local circumstances and feedback. This interesting empirical issue is still under investigation. Although research suggests that people adapt their heuristics to the local situation and to performance feedback through a process of reinforcement learning (Rieskamp and Otto, 2006), that does not prevent heuristics from overshooting or being misapplied.

[8] Standard examples of mismatch include our craving for fat and sugar (leading to widespread obesity), or our disproportionate fear of spiders (given that very few people succumb to spider bites), compared to our carelessness about moving cars and electricity sockets (which are far more hazardous).

[9] It is worth pointing out that the hot hand hypothesis is not per se implausible, and certainly would not overthrow our current understanding of the world (as with the gambler's fallacy). Maybe players who start a game well get a boost of confidence, and those who start out badly experience more stress, which further lowers their performance.

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

Are these unflattering labels justified? Steven Pinker has argued that the gambler's fallacy is not really a fallacy, because in most natural environments, a succession of events is not statistically independent: the probability of an event changes in function of earlier occurrences.

> Many events work like that. They have a characteristic life history, a changing probability of occurring over time which statisticians call a hazard function. An astute observer should commit the gambler's fallacy and try to predict the next occurrence of an event from its history so far ...
>
> [Pinker, 1997, p. 346]

Roulette wheels, with their smooth and radially symmetric design, are expressly designed to foil this heuristic, according to Pinker: 'in any world without casinos, the gambler's fallacy is rarely a fallacy' (Pinker, 1997). Indeed, our ability to spot statistical dependencies between events in real-life may often prove very useful. That morale even holds for chance games outside the idealized world of casinos. If I play a dice game at a fair and the dice land on 12 a couple of times in a row, it is not unreasonable to predict that they will do so the next time too – because they are probably rigged. In general, as Bennis et al. have written, 'casino games [are] exquisitely designed to exploit otherwise adaptive heuristics to the casino's advantage' (Bennis, Katsikopoulos, Goldstein, Dieckmann, & Berg, 2009, p. 421). Taleb (2008) has referred to the inappropriate application of pure and simplified models of probability to real life as the 'ludic fallacy'. Intuitions that lead us astray in the rarefied and idealized world of casinos, may prove very useful in real-life situations.

If I am presented with some strange device that churns out black and white balls, and I do not have access to its inner working, I may be forgiven for thinking that there is a statistical dependence between different trials. I may even be perfectly justified in doing so, extrapolating from earlier experience. An experienced gambler in a casino, by contrast, possesses all the requisite evidence he needs to conclude that this is not a machine to which his pattern detection heuristics will apply. If even a crash course in statistics and a careful inspection of the roulette wheel fail to cure him of his habit of thought, then we may, *pace* Pinker, call his reasoning fallacious.

The natural way to construe Pinker's argument is as a point about adaptation. Given that pure randomness is rare in the EEA (as it most probably still is today in natural environments), it made sense for evolution to endow us with a knack for spotting patterns, with the minor side-effect of making us vulnerable to (fair) dice games and roulette wheels. After enduring five straight days of rainy weather, as Pinker notices, it would have made sense for the optimistic early hominid to expect a sunny spell. As weather fluctuation often follows statistical patterns, the sun can really be 'overdue'. For evolution, it makes sense to 'gamble' on the kind of environments with a specific hazard function, and to disregard the abstractions of casino wheels, which did not feature in ancestral environments. In a similar vein, the hot hand fallacy may arise from a heuristic adapted to environments with positive temporal autocorrelation or 'clumping' (Fawcett et al., 2014; Wilke & Barrett, 2009).

But what does that say about us? Are we forgiven for using the heuristic anyway, even in the face of countervailing information? Personal-level rationality involves the ability to use contingent information about the problem task in order to understand which strategies are conducive to success. A rational person should reflect upon the deliverances of her intuitions and rules of thumb, and disregard them when she has good reasons to believe that they lead her astray (Over, 2000). Irrationality often stems from what Fiedler and Wänke (2004, 2013) have called 'meta-cognitive myopia', or the inability to critically evaluate intuitive heuristics and use them in appropriate settings. In the case of casinos, there is indeed an alternative course of action available. Many experienced gamblers, though still feeling the intuitive pull of the pattern-seeking heuristic, suppress or resist it in practice, because they realize that the wheel has no memory and every turn is independent of all the others. Surely (Dennett, 2013) such players are more 'rational' than the ones who succumb to the gambler's fallacy, in all of the senses we explored: they understand the justification of their beliefs, they intentionally aim for a specific goal, and they will be more successful in attaining it. Indeed, one of the hallmarks of (human) rationality is our ability to reflectively evaluate and cross-validate the output of our cognitive modules, and to overrule their intuitive hunches when the context calls for it. In this regard, whereas Pinker emphasises that the gambler's fallacy is not a fallacy in a world without casinos, we emphasise that it is nevertheless a fallacy in a world *with* casinos.

As de Sousa points out, the rational 'strategies' attributed to natural selection are conceived in the most general terms, on the basis of phylogenetic 'experiences' and without precise reference to the circumstances under which they may be implemented. By contrast, individual rational agents face particular problems in possibly unique circumstances (de Sousa, 2007, p. 137). Rationality, in this perspective, consists of using our reasoning abilities appropriately to deal with the situation at hand, not blindly following heuristics of which – with hindsight – we can appreciate the adaptive rationale.

5. Ecological rationality restored

5.1. Conjunction fallacy?

Does the program of ecological rationality always involve locus shifts, and their attendant problems of transparency and intentionality? No. The second strand of the program of "ecological rationality" (Section 2.2), does not depend on adaptive considerations, and hence does not involve locus shifts. The 'ecological' dimension of rationality, in this strand, concerns the real-life contexts of human reasoning, which are typically truncated in artificial lab experiments. To illustrate this difference, we will briefly consider two examples of alleged human bias that have been reinterpreted under the banner of "ecological rationality", but which do not involve adaptive locus shifts and which satisfy the minimal requirements of personal-level rationality. Consider the following classic test:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.
>
> Which is more probable?

(A) Linda is a bank teller.
(B) Linda is a bank teller and is active in the feminist movement.

In their original research on what became known as the 'Linda problem', Kahneman & Tversky held human reasoners accountable to the conjunction rule: B cannot be more probable than A, because it is an accepted rule of probability theory that the conjunction of two events can never be more probable than either of its members. Still, the majority of subjects answered B.

Gigerenzer, however, pointed out that terms such as 'and' and 'probable' are polysemous. 'Probable' can also mean plausible, sensible, or supported by evidence (Gigerenzer, 1996; Hertwig & Gigerenzer, 1999). Kahneman & Tversky expect subjects to interpret 'probable' in the sense of mathematical probability. As a number of researchers have pointed out, however (Adler, 1984; Dulany & Hilton, 1991; Hilton, 1995), this construal violates pragmatic rules of conversational inference, in particular the maxim of relevance (Grice, 1989). According to this maxim, it makes little sense to adopt the notion of mathematical probability in answering the Linda problem, given all the additional (and hence presumably relevant) information presented. If subjects interpret the task as Kahneman & Tversky want them to, the little vignette about Linda is rendered completely irrelevant to the task, and the 'correct' answer depends solely on the logical operator 'AND' and the mathematical meaning of 'probable'. Why would the researcher present such an elaborate description of Linda's character and background, if it were completely irrelevant? In addition, subjects can reasonably interpret the first statement as implying that Linda is *not* active in the feminist movement. These ambiguities can be ruled out by presenting the task in terms of natural frequencies, or by including another question that makes the description of Linda relevant, so that the (overall) task does not violate the maxim of relevance. After such modifications, as it turns out, subjects reason in accordance with the conjunction rule (Hertwig & Gigerenzer, 1999).[10] Given that simple application of the conjunction rule glosses over the conversational subtleties and ambiguities of the situation, a number of critics have argued that Kahneman & Tversky's interpretation is uncharitable (e.g. Margolis, 1987), and even that *they*, and not their subjects, have failed to grasp the logic of the Linda problem (Hertwig & Gigerenzer, 1999): 'human intelligence reaches far beyond narrow logical norms. In fact, the conjunction problems become trivial and devoid of intellectual challenge when people finally realize that they are intended as a content-free logical exercise' (Gigerenzer, 2008, p. 73).

5.2. Base rate fallacy?

Another case of alleged irrationality that (at least partly) disappears when cast in an appropriate real world environment, instead of an artificial and misleading set-up, is the infamous base rate fallacy (Casscells, Schoenberger, & Graboys, 1978; Kahneman et al., 1982, p. 154). For example, when estimating the probability that a patient has contracted a particular disease, physicians are (sometimes) found to ignore the base rate of the disease in question. Information about the reliability of the test appears to be the sole determinant of their judgment, while the statistical prevalence of the disease in the population group is ignored.

However, as Gigerenzer noted, real life is more complicated

> Clinicians, however, know that patients are usually not randomly selected—except in large survey studies—but rather 'select' themselves by exhibiting symptoms of the disease. In the absence of random sampling, it is unclear what to do with the base rates specified.
>
> [Gigerenzer, 1991, p. 9]

If patients are not randomly sampled, Gigerenzer notes, physicians cannot be accused of committing the base rate fallacy. Indeed, given that real world environments, as opposed to convenient mathematical idealizations, do not typically proffer random samples, the judgment of the clinicians is perfectly defensible. In a series of experiments conducted by Gigerenzer and et al. (1988), people were found to be capable of factoring in base rates, provided they were 'convinced' that the sample was randomly drawn from the population. Inviting subjects to blindly draw patients' names from an urn proved effective in overruling their prior beliefs, whereas merely pointing out random sampling in the problem description did not.

Even though non-random sampling appears to be an intuitive default assumption, these experiments show that people can and will – given sufficient commitment – account for base rates when necessary. Indeed, as Gigerenzer points out, with

[10] However, in Kahneman's defense, other experiments show that, even if the Linda problem is presented in a frequency format, which is reputedly ecologically valid, biased reasoning does sometimes – though not always – persist. (Mellers, Hertwig, & Kahneman, 2001).

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025

Tooby and Cosmides (1994), the neglect of base rate information is by no means a systematic defect of human reasoning, as the label 'base rate fallacy' suggests. If Bayesian problems are couched in terms of frequencies, as opposed to abstract mathematical probability, it turns out that the 'fallacy' largely disappears (but see Kahneman & Tversky, 1996; Sloman, Over, Slovak, & Stibel, 2003). In other words, it is not that most people lack the requisite skills to integrate base rate information (Zhu & Gigerenzer, 2006). Provided that there is sufficiently strong indication of random sampling, and that base rate information is presented in a concrete, accessible way, people's reasoning approximates Bayesian theory (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Once again, Gigerenzer and colleagues argue that the 'fallacy' is elicited by the artificial or experimental set-ups, and evaporates when transferred to real life. It is important to note that whether and to what extent the conjunction and base rate fallacies disappear when problems are presented in ecologically valid formats is still hotly debated. This, however, is a matter of experimental psychology. The point we are making is merely that the approach taken in these cases does not involve a problematic locus shift.

5.3. Analysis

Wherein lies the difference between the latter cases and the problematic ones we discussed earlier, given that both involve a form of 'ecological' reframing? In Gigerenzer's deflation of the 'conjunction fallacy' or the 'base rate fallacy', the gestalt switch also involves adopting a broader view, casting human inference-making in a rich and real world environment. In these examples, however, the force of the ecological gestalt switch derives not from any shift to the adaptive rationale of their heuristics, but from the fact that human reasoners had not been given sufficient information to disambiguate the problem at hand and to home in on the interpretation intended by the experiments. In particular, subjects did not know – or were not properly committed to the belief – that they were dealing with an artificial situation in which the normal richness and ecological complexity of problem solving should be ignored. How were Kahneman & Tversky's subjects supposed to know that their experimenters were interested in a 'silly' question that violates an accepted rule of sensible conversation? Why *should* they have interpreted the problem in the sense of mathematical probability, contra their intuition that this made the whole story about Linda irrelevant? Similarly, why should the physicians have realized they were expected to take a step back from their regular practice, considering an unrealistic case in which a subject is randomly drawn from the population at large. Remember that, when Gigerenzer (1991) committed subjects to the mathematical interpretation, their reasoning conformed to the conjunction rule. The force of the argument from ecological rationality here is not that 'we can design experiments in which cognitive illusions disappear' (Kahneman, 2003, p. 711), as Kahneman construes Gigerenzer's point, but that *there was no illusion in the first place*.

6. Conclusion

In order to understand the strengths and frailties of human reason within a naturalized framework, an evolutionary perspective is invaluable. We applaud the program of ecological rationality for taking the evolutionary roots of (human) cognition seriously. If one tries to dispel apparent instances of human irrationality, however, one should not take recourse to ultimate, evolutionary alibis. The foibles of reason cannot be exculpated simply because they display some evolutionary rationale, whether arising from error management, adaptive bias, or some mismatch between evolved heuristics and modern environments. It is no criticism of a screwdriver (or its human designer) to note that it makes a poor crowbar – but this does not exculpate the DIY enthusiast who uses a screwdriver as a lever. Likewise, we can admire the adaptive or functional design of a heuristic while still impugning the rationality of individuals who blindly misapply it – and should know better.

Conversely, evidence of adaptive cognitive design is not evidence of human rationality. We cannot fully congratulate ourselves on our 'rational' behaviour unless we have some cognitive access to the goal we want to achieve, our strategy for attaining it, and our reasons for thinking this strategy is or might be successful. Insofar as humans are working on automatic pilot, profiting from (or being misled by) the wisdom of their evolved heuristics, the normative categories of (ir)rationality do not apply. Genuine rationality presupposes an intentional striving for success.

The program of ecological rationality has pursued two distinct projects, although the distinction between these has been underappreciated. In its best moments, it has convincingly demonstrated that the rigid application of content-blind logical norms as a benchmark of rationality produces an uncharitable view of human cognition, a view insensitive to the subtlety and richness of human reasoning. On the other hand, advocates of ecological rationality, together with evolutionary psychologists, have also pointed out the adaptive rationale of some of our cognitive biases. That research is certainly fascinating, but we should remain wary of subtle locus shifts between different levels of rationality. In some cases (e.g. superstition, gambler's fallacy), having an adapted mind is compatible with misbelief and irrationality. Even sheer stupidity.

Acknowledgments

The research of the first author was supported by the Fund for Scientific Research Flanders (FWO), and carried out at the Konrad Lorenz Institute in Vienna. We are very grateful to both the editors and the reviewers for their helpful and thoughtful feedback.

References

Adler, J. E. (1984). Abstraction is uncooperative. *Journal for the Theory of Social Behaviour, 14*(2), 165–181.

Aktipis, C. A., & Kurzban, R. (2004). Is Homo economicus extinct?: Vernon Smith, Daniel Kahneman and the evolutionary perspective. In R. Koppl (Ed.), Advances in Austrian economics (pp. 135–153).

Ariely, D. (2009). *Predictably irrational, revised and expanded edition: The hidden forces that shape our decisions*. HarperCollins.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486–498.

Babler, T. G., & Dannemiller, J. L. (1993). Role of image acceleration in judging landing location of free-falling projectiles. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 15.

Bennis, W., Katsikopoulos, K., Goldstein, D., Dieckmann, A., & Berg, N. (2009). Designed to fit minds: Institutions and ecological rationality. In G. Gigerenzer & P. M. Todd (Eds.), *Ecological rationality: Intelligence in the world* (pp. 409–427). New York: Oxford University Press.

Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *The New England Journal of Medicine, 299*(18), 999.

Cosmides, L., & Tooby, J. (1994). Better than rational: Evolutionary psychology and the invisible hand. *The American Economic Review, 84*(2), 327–332.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*(1), 1–73.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press, Incorporated.

de Sousa, R. (2007). *Why think? : Evolution and the rational mind: Evolution and the rational mind*. New York: Oxford University Press, USA.

de Sousa, R. (2009). Evolution, thinking, and rationality. In M. Ruse (Ed.), *Philosophy after Darwin: Classic and contemporary readings* (pp. 289–300). Princeton, NJ: Princeton University Press.

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences, 6*(03), 343–355.

Dennett, D. C. (1995). *Darwin's dangerous idea*. London: Penguin.

Dennett, D. C. (1996). *Kinds of minds towards an understanding of consciousness*. London: Weidenfeld & Nicolson.

Dennett, D. C. (2013). *Intuition pumps and other tools for thinking*. WW Norton & Company.

Dulany, D. E., & Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition, 9*(1), 85–110.

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., et al (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*.

Fiedler, K., & Wänke, M. (2004). On the vicissitudes of cultural and evolutionary approaches to social cognition: The case of meta-cognitive myopia. *Journal of Cultural and Evolutionary Psychology, 2*(1), 23–42.

Fiedler, K., & Wänke, M. (2013). Why simple heuristics make life both easier and harder: A social-psychological perspective. In R. Hertwig & U. Hoffrage (Eds.), *Simple heuristics in a social world* (pp. 487–515). New York: Oxford University Press.

Galperin, A., & Haselton, M. G. (2012). Error management and the evolution of cognitive bias. In J. P. Forgas, Sedikedes Fiedler, & C. (Eds.), *Social thinking and interpersonal behavior* (pp. 45–64). New York: Psychology Press.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of Social Psychology, 2*(1), 83–115.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*(3), 592–596.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.

Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford: Oxford University Press.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107–143.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 513.

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684–704.

Gigerenzer, G., & Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox*. MIT Press.

Gigerenzer, G., & Sturm, T. (2012). How (far) can rationality be naturalized? *Synthese, 187*(1), 243–268.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology, 44*(6), 1110–1126.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases. The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*(3), 295–314.

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*(1), 75–90.

Grice, P. (1989). *Studies in the way of words*. Harvard: Harvard University Press.

Hamilton, W. D. (1963). Evolution of altruistic behavior. *American Naturalist, 97*(896), 354–356.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78*(1), 81–91.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*(1), 47–66.

Haselton, M. G., Nettle, D., & Andrews, P. W. (2005). The evolution of cognitive bias. *The Handbook of Evolutionary Psychology*, 724–746.

Hertwig, R., & Gigerenzer, G. (1999). The conjunction fallacy revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275–306.

Hertwig, R., & Hoffrage, U. (2013). *Simple heuristics in a social world*. New York: Oxford University Press.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin, 118*(2), 248–271.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*(3), 326.

Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature, 477*(7364), 317–320.

Kacelnik, A. (2006). Meanings of rationality. In S. L. Hurley & M. Nudds (Eds.), *Rational animals?* (pp 87–106). Oxford: Oxford University Press.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697–720.

Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin Books.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*(3), 582–591.

Kihlstrom, J. F. (2004). Is there a "People are Stupid" school in social psychology? [commentary on "towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition" by J.I. Krueger and D.C. Funder]. *Behavioral & Brain Sciences, 27*, 348.

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27*(3), 313–327.

Lorenz, K. (1941). Kant's doctrine of the a priori in the light of contemporary biology. In M. Ruse (Ed.), *Philosophy after Darwin. Classic and contemporary readings* (pp. 231–247). Princeton, NJ: Princeton University Press.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review, 118*(3), 393.

Margolis, H. (1987). *Patterns, thinking and cognition a theory of judgment*. Chicago London: University of Chicago press.

Marshall, J. A. R., Trimmer, P. C., Houston, A. I., & McNamara, J. M. (2013). On evolutionary explanations of cognitive biases. *Trends in Ecology & Evolution, 28*(8), 469–473.
Maynard Smith, J. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology, 47*(1), 209–221.
McKay, R., & Efferson, C. (2010). The subtleties of error management. *Evolution and human behavior, 31*(5), 309–319.
Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12*(4), 269–275.
Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57.
Millikan, R. G. (1984). *Language thought and other biological categories: New foundation for realism*. MIT Press.
Nesse, R. M. (2001). The smoke detector principle. *Annals of the New York Academy of Sciences, 935*, 75–85.
Over, D. (2000). Book review: Ecological rationality and its heuristics. *Thinking & Reasoning, 6*(2), 182–192.
Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). When is the recognition heuristic an adaptive tool? In P. M. Todd & G. Gigerenzer (Eds.), *Ecological rationality: Intelligence in the world* (pp. 113–143). New York: Oxford University Press.
Piattelli-Palmarini, M. (1996). *Inevitable Illusions: How Mistakes of Reason Rule Our Minds*. Wiley.
Pinker, S. (1997). *How the mind works*. New York: Norton.
Raab, M., Gula, B., & Gigerenzer, G. (2012). The hot hand exists in volleyball and is used for allocation decisions. *Journal of Experimental Psychology: Applied, 18*(1), 81–94.
Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236.
Sharot, T. (2011). The optimism bias. *Current Biology, 21*(23), R941–R945.
Shermer, M. (2011). *The believing brain: From ghosts and gods to politics and conspiracies – How we construct beliefs and reinforce them as truths*. Henry Holt and Co..
Singer, B., & Benassi, V. A. (1981). Occult beliefs: Media distortions, social uncertainty, and deficiencies of human reasoning seem to be at the basis of occult beliefs. *American Scientist, 69*(1), 49–55.
Skyrms, B. (2000). Game theory, rationality and evolution of the social contract. *Journal of Consciousness Studies, 7*(1–2), 1–2.
Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91*(2), 296–309.
Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Cambridge, Mass.: Blackwell.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645–665.
Stanovich, K. E., & West, R. F. (2003). Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. *Evolution and the psychology of thinking: The debate*, 171–230.
Sutherland, S. (2007). *Irrationality*. Ebbw Vale, Wales: Pinter & Martin.
Taleb, N. N. (2008). *The black swan: The impact of the highly improbable*. London: Penguin Books Limited.
Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
Tooby, J., & Cosmides, L. (1994). The biological foundations of culture. In J. Barkow, J. Tooby, & L. Cosmides (Eds.), *The adapted mind. Evolutionary psychology and the generation of culture* (pp. 19–136). New York, New York, USA: Oxford University Press.
Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evolution and Human Behavior, 30*(3), 161–169.
Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton (N.J.): Princeton University Press.
Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition, 98*(3), 287–308.

Please cite this article in press as: Boudry, M., et al. Can evolution get us off the hook? Evaluating the ecological defence of human rationality. *Consciousness and Cognition* (2014), http://dx.doi.org/10.1016/j.concog.2014.08.025