Reliabilist Epistemology Meets Bounded Rationality

Giovanni Dusi

Autonomous University of Barcelona

Giovanni.Dusi@uab.cat

Abstract

Epistemic reliabilism holds that a belief is justified if and only if it is produced by a reliable or truth-conducive process. I argue that reliabilism offers an epistemology for bounded rationality. This latter concept refers to normative and descriptive accounts of real-world reasoning instead of some ideal reasoning. However, as initially formulated, reliabilism involves an absolute, context-independent assessment of rationality that does not do justice to the fact that several processes are reliable in some reasoning environments but not in others, as is widely reported in the cognitive sciences literature. I consider possible solutions to this problem. Resorting to 'normality reliabilism', a variant of the theory, is one; but I find it insufficient. Therefore, in addition, I propose to relativise the reliability assessment to reasoning environments. This novel version of reliabilism fits bounded rationality better than the original one does.

Keywords

Process reliabilism – Ecological rationality – Heuristics and biases – Normality reliabilism – Heuristics – Reasoning environments

1. Introduction

According to reliabilism, the epistemic theory first introduced by Alvin Goldman, a belief is justified if and only if it is produced by a reliable belief-forming process, which is one that usually generates true beliefs rather than false ones. In this essay, I argue that reliabilism offers an epistemology for bounded rationality, a term used to refer to a broad spectrum of normative and descriptive accounts of real-world reasoning instead of some ideal reasoning. However, my argument continues, reliabilism can only provide this service as long as it is subject to changes that lend a more contextual character to the judgement of reliability. Indeed, no version of the theory, neither the original one nor those proposed later, can account for reasoning processes that are reliable in certain types of reasoning environments and unreliable in others, while that this does occur is an empirical fact well-known in cognitive sciences.

In my discussion, I will assume that there is a close connection between epistemological investigation and empirical sciences. This naturalistic approach to epistemology has been adopted in several philosophical inquiries into justification and rationality conducted taking into consideration the findings of the psychology of reasoning (for instance, Goldman, 1986, 2008; Bishop & Trout, 2005; Kornblith, 2014; Sturm, 2020). The idea is that the psychological sciences offer empirical results about how we actually think and reason that might be useful or essential to make progress in addressing epistemic questions (cf. Feldman's [2012] 'cooperative naturalism').

It should be noted that reliabilist epistemology and bounded rationality have already met. The venue was a colloquium on the foundation of the norms of rationality between leading philosophers and psychologists[1]. On that occasion, Alvin Goldman gave a talk on epistemological and psychological perspectives on human rationality. In his article which followed the colloquium, Goldman acknowledged that: "In fact, each of the other contributors to this symposium, Gerd Gigerenzer and Michael Bishop, has made roughly this proposal; each regards truth, or accuracy, as (at least) one type of consequence that is pertinent to a rationality appraisal." (2008, p. 240). Gerd Gigerenzer is a cognitive psychologist who has propounded an original approach to bounded rationality called 'ecological rationality'. His reply to Goldman arrived later in a 2019 article, where he claimed that "ecological rationality can… be evaluated by goals beyond coherence, such as predictive accuracy, frugality, and

[1] More precisely, the colloquium was part of the Sixth International Conference organised by the German Society for Analytic Philosophy (GAP) which took place in 2006 in Berlin.

efficiency" (p. 3547) and that: "The extension of goals of rationality from coherence to performance… has much in common with existing approaches, such as Kitcher's (1992) naturalism and Goldman's (1999) epistemological reliabilism." (p. 3556).

The take-home message from this exchange over the years seems to be that process reliabilism (*PR*, henceforth) fits Gigerenzer's ecological approach to bounded rationality (*EBR*, from now on) pretty well. Assessing this match is my primary goal here. To make this assessment possible, I will need to go into the details of both standpoints. Thus, in §2, I introduce bounded rationality, focusing on *EBR*. I then (§3) present reliabilism and Goldman's paradigmatic examples of reliable and unreliable processes. A relevant difference between *PR* and *EBR* will manifest itself at that stage. In §4, I will show how Goldman's original, simple version of reliabilism involves an absolute assessment of rationality, which does not do justice to reasoning processes that are reliable in some contexts but not in others. This is problematic since the psychological sciences have shown how the performance of several reasoning strategies changes from one environment to another. In §5, I start searching for solutions to this problem which preserve the match between *PR* and *EBR* by appealing to a variant of *PR* called *normality reliabilism* or *normal conditions reliabilism*. Despite specifying 'normal' and 'abnormal' domains for the reliability assessment, I will claim that normality reliabilism fails to solve the problem of absolute reliability. Even so, the normality condition proves helpful to block some of the standard objections to reliabilism, so we should keep it rather than drop it. In the following section, I advance my proposal to relativise the reliability assessment to reasoning environments. The novel version of *PR* that will emerge can account for differences in the performance of one and the same process between one reasoning environment and another. Thus, my reliabilist epistemology offers a theory of rational belief that fits bounded rationality better than the original version does. In §7, I draw some conclusions.

2. Bounded rationality

The conditions in which human and non-human agents find themselves when drawing inferences or making decisions are often less than ideal. For instance, agents could evaluate alternative options having only scarce information about the alternatives. Similarly, they could need to forecast uncertain future events that are beyond easy prediction or lack the time needed to consider all relevant factors and be forced to come to an answer quickly. Moreover, both animal and machine computational capabilities have intrinsic limitations. For example, the storage capacity of human working memory is restricted to a small number of meaningful pieces of information. For these and similar reasons, the term 'bounded rationality' was introduced to qualify the reasoning capacities of agents who draw inferences and make decisions under environmental and cognitive constraints. Economist Herbert Simon initially proposed it as a departure from the conception of rationality common in neoclassical economics (Simon, 1957). Since then, it has been used in numerous other fields, including computer science, decision sciences, cognitive psychology, neuropsychology, biology, and philosophy.

Research on bounded rationality encompasses several accounts of behaviour which deviates from exhibiting ideal rationality. To begin with, it is possible to distinguish between descriptive and normative theories. Moreover, specific instances of these two kinds of theories are sometimes combined in distinct views on human reasoning. In the following two subsections, I briefly present the main accounts of such behaviour and then focus on *EBR*, one specific view.

2.1. Normative and descriptive accounts

Part of the inquiry concerning bounded rationality addresses the normative standards against which judgements and decisions are evaluated. These standards are those people are expected to reason by. Broadly speaking, there are two main views. The first one, the *coherence-based* view, takes compliance with rules of logic, probability, and decision theory as the rationality criterion. Dubbed by Stein (1996) the 'standard picture of rationality', this view considers reasoning to be at fault whenever it does not comply with these rules (cf., for instance, Piattelli-Palmarini, 1994).

Alternatively, the *process-based* view understands rationality as the use of appropriate reasoning processes (cf., for instance, Simon, 1976). The appropriateness of processes is often evaluated in consequentialist ways, such as promoting one or more epistemic values. Accuracy—the production of true beliefs while not producing false ones—is the value most often mentioned in this regard. Process-based theories of rationality that prioritise accuracy hold that reasoning is rational if it employs processes that allow it to arrive at the correct output in a cognitive task. When the rationality criterion is that it applies accurate processes, reasoning is evaluated on its performance.

Regarding empirical research, cognitive psychology is the discipline that has incited most debate about the bounds of reasoning. Here, it is possible to recognise two dominant lines of thought. The first trend is associated with the *Heuristics and Biases* programme (*H&B*, henceforth; see Kahneman et al., 1982; Gilovich et al., 2002). According to this view, reasoners' cognitive limitations make it challenging to reason according to the rules of logic, probability, and decision theory. Thus, people frequently use heuristics to reduce complex tasks, such as assessing probabilities and predicting values, to simpler judgemental operations. However, heuristics often depart from the standard norms and tend to make people fall into systematic reasoning mistakes called biases. For example, people are prone to commit the conjunction fallacy when they evaluate the probability of two co-occurring events (Tversky & Kahneman, 1983), or they tend to neglect relevant base-rate information when predicting uncertain events (Kahneman & Tversky, 1973).

The *Ecological Rationality* programme promotes the second prominent view in this debate (see Todd & Gigerenzer, 2012; Gigerenzer et at., 1999). Supporters of this view have drawn attention to the capacity of animal and artificial agents to take advantage of environmental structures to draw intelligent inferences and make smart decisions. People can adapt their cognitive strategies to the context in which reasoning occurs, using those strategies that fit the features of the environment and result in good reasoning. Most notably, it has been empirically demonstrated that simple heuristics can enable people to draw correct inferences, and sometimes even more correct inferences than those produced by complex strategies that obey logical and probabilistic norms (see Gigerenzer et al., 2011). For instance, simple recognition heuristic can outperform Wimbledon experts and ATP rankings in predicting the outcomes of Wimbledon matches (Serwe & Frings, 2006); and fast and frugal decision trees can be more accurate than logistic regressions and other complex statistical models in several tasks, such as predicting bank failure (Aikman et al., 2014) or in medical diagnoses such as detecting depressed mood (Jenny et al., 2013).

Some combinations of normative and descriptive accounts might have already become evident to the reader. On the one hand, *H&B* subscribes to the coherence-based theory of rationality and evaluates reasoning on its respect for standard rules. On the other hand, *EBR* assesses as rational those inferences and decisions which fulfil the epistemic goals, embracing the process-based and consequentialist standpoint (cf., for instance, Schurz & Hertwig, 2019).

The differences between the two schools of thought are substantial. They represent two cultures of research on bounded rationality, telling two different stories about rationality (Katsikopoulos, 2014). These differences fostered debate within the psychology domain in the 1990s, sometimes leading to harsh clashes[2]. The two psychology programmes portray images of human rationality in different shades of colour: *H&B* shows numerous ways humans fail to conform to the standard rules of reasoning, while *EBR* points out many instances of adaptive and successful behaviour.

So far, I have distinguished between normative and descriptive accounts of bounded rationality and illustrated how they can be combined in existing research programmes. I will now focus on *EBR*, given its compatibility with epistemic reliabilism.

2.2. The ecological approach to bounded rationality

EBR is committed to explaining how animal and artificial agents cope with the bounds of reasoning. The empirical research conducted within this programme encompasses cases of both

[2] The debate between Tversky and Kahneman of *H&B* on one side and Gigerenzer of *EBR* on the other gave rise to a considerable body of literature. The discussion started from an adversarial exchange which has become a classic in the field of theories of rationality (see Gigerenzer, 1991; Kahneman & Tversky, 1996; Gigerenzer, 1996). In philosophy, several authors have attempted to frame and tame the disagreement. See Vranas (2000), Samuels et al. (2002), and Sturm (2012, 2019) for alternative ways to address the issue.

theoretical and practical reasoning. For example, it studies how reasoners form beliefs when deciding between two alternatives (e.g., which city has more inhabitants, Milan or Brescia?) and how agents choose the course of action to take (e.g., how should money be invested in a certain number of assets?). However, here I will only look at cases of theoretical reasoning and consider bounded rationality only as a form of epistemic rationality[3].

According to the ecological approach, every agent facing an inferential or decisional task has recourse to a plurality of reasoning strategies for addressing the task. These are algorithmic procedures of different kinds. Some are complex and make extensive calculations; others, the heuristics, are fast and frugal processes that produce outcomes in little time and using few pieces of information. Altogether, they constitute the 'cognitive toolbox', the repertoire of tools for drawing inferences and making decisions. It is crucial that the reasoner selects the strategy that best suits the reasoning task and the context in which the cognitive performance takes place. Indeed, using the appropriate reasoning method substantially contributes to reasoning success (cf. Gigerenzer & Sturm, 2012, pp. 251-252; Hertwig et al., 2021, pp. 17-18).

Ecological rationality builds on Herbert Simon's adaptive view of rational behaviour. As Simon puts it, "Human rational behavior... is shaped by a scissors [sic] whose two blades are the structure of the task environments and the computational capabilities of the actor" (Simon, 1990, p. 7). This claim does not have to be understood in descriptive terms only. Indeed, it contains an important normative message: what is to be considered rational behaviour depends on the relationship between the reasoning environment and the cognitive capacities of the reasoner. Ecological rationality emerges when the

[3] These choices need to be justified. The reason for the first deals with the scope of my research. I aim to delineate an account of bounded *epistemic* rationality. In other words, I want to examine what it means for reasoning to be epistemically rational under constraints of information, time, cognitive resources, etc. Moreover, the instances of bounded reasoning I am interested in can be evaluated on their epistemic merits without considering any practical virtues. This means it is possible to assess their normative status without referring to practical issues; hence the second choice.

structure of reasoning mechanisms matches the structure of the environment and the cognitive setup of the reasoner. Therefore, no judgement of rationality can be made without reference to the context in which the reasoning is performed.

Ecological rationality, normatively understood, is about the success of reasoning strategies in the world, assessed by the following principle (Brighton & Todd, 2009, p. 337):

> ECOLOGICAL RATIONALITY
>
> A mechanism M is ecologically rational in environment E in comparison to some other mechanism M' when M outperforms M' on some criterion, or currency, of comparison.

The criterion of comparison most often used is the accuracy of judgement. However, other criteria are also sometimes adopted, such as speed (how fast an inference can be drawn or a decision made), frugality (how many pieces of information need to be searched before a judgement can be made), etc.

It might be helpful to see an example from the *EBR* literature to understand how this principle works. Consider this inferential task:

> *Creditworthiness*
>
> of two companies, you have to choose the more creditworthy.

To draw the appropriate inference[4], agents usually use information gathered from different cues such as the company's financial flexibility, its efficiency, the qualifications of the employees, etc. Psychological research has shown that these cues can be processed in somewhat different ways

[4] Someone might see this as a case of decision-making rather than belief formation. However, this can *also* be a case of belief formation, as one might think that when I face such a reasoning task, in addition to choosing between two alternatives, I also form a belief as to which company is more creditworthy. This is the sense in which I am interested in this case.

(Gigerenzer & Goldstein, 1996). Suppose all the cues are binary (0 and 1), where 1 is the higher criterion value, and they can be ranked according to their validity. By cue validity, we mean the probability of drawing a correct inference on the condition that the cue used in drawing that inference discriminates between the options. The higher the cue validity, the higher the probability of drawing a correct inference when relying on that cue. In our example, this is the probability of correctly inferring which of two companies is more creditworthy relying on a cue whose value is positive (1) for one company and non-positive (0 or unknown) for the other[5]. The first strategy people might choose from their cognitive toolbox is the linear weighted additive strategy (WADD). For each company, WADD calculates the sum of all cue values multiplied by the cue's validity and then selects the alternative with the largest sum. Alternatively, there is evidence that people rely on a simpler method known as the take-the-best heuristic (TTB). TTB is a one-reason decision-making method that decides which of two objects scores higher on a criterion solely on the basis of the most valid cue that discriminates between the two objects. This strategy consists of three steps (Gigerenzer, 2008, p. 32):

1. *Search rule*: Search through cues in order of their validity. Look up the cue values of the cue with the highest validity first.

2. *Stopping rule*: If one object has a positive cue value and the other does not (or is unknown), stop the search and proceed to Step 3. Otherwise, ignore this cue and return to Step 1. If no more cues are found, guess.

3. *Decision rule*: Predict that the object with the positive cue value has the higher value on the criterion.

[5] More precisely, the validity v of a cue is estimated as:

$$v = C / (C + W),$$

where C is the number of correct inferences when a cue discriminates, and W is the number of wrong inferences, all estimated from samples (see Gigerenzer, 2019, p. 3559).

How each strategy performs is a matter of empirical research, and the evidence gathered by proponents of *EBR* shows something quite surprising. Despite its simplicity, TTB matches alternative, more complex and cognitively demanding strategies, including WADD, in inferential accuracy. This happens when specific circumstances obtain, for example, when the information available to draw the inference is *non-compensatory*. This concept refers to the fact that each cue available for choosing between the two alternatives cannot be outweighed by any combination of less valid cues. To see what this means, consider the following explanation from Gigerenzer (2008, p. 37). Assume an environment with *M* binary cues ordered according to their weights W_j, with $1 \leq j \leq M$. A set of cue weights is non-compensatory if the following property holds for each cue's weight:

$$W_j > \sum_{k>j} W_k$$

As an example, take a set of weights that decreases exponentially, such as 1, 1/2, 1/4, 1/8, and so on. The sum of the cue weights to the right of a cue can never be larger than this cue's own weight. In environments with this feature, namely, non-compensatory environments, no linear strategy can outperform the faster and more frugal TTB, achieving at most the same level of accuracy as this simple heuristic[6] (Martignon & Hoffrage, 2002).

Furthermore, TTB can also outperform WADD when the non-compensatory condition holds (Brighton & Gigerenzer, 2015). This phenomenon is called the *less-can-be-more* effect and is one of the most astonishing findings of *EBR* research[7]. Under non-compensatory conditions, using TTB is

[6] Non-compensatoriness is an environmental feature prevalent in natural environments to such a high degree that inferences guided by this feature approach (and occasionally exceed) the inferences drawn using multiple linear regression in predictive accuracy (Şimşek, 2013).

[7] The reason why a simple method can be more accurate than complex ones deals with the nature of the predictive error. The total error of predictive models is, roughly speaking, composed of two parts: bias and variance. When cues and cue values have the properties we assumed here, and the information is non-compensatory, TTB and WADD have the same bias, but TTB tends to have a smaller variance. Hence, TTB's total error can be less than that of WADD (for further details, see Wheeler, 2020).

ecologically rational. Alternatively, if this does not hold, and for instance we have compensatory conditions, then TTB performs poorly and is usually outperformed by WADD. TTB is not ecologically rational in compensatory environments, and people would be better off relying on other strategies, such as WADD.

This example shows how rationality is assessed in the ecological framework. Each of the two strategies for solving *Creditworthiness* is more accurate than the other in certain environments. Therefore, given the definition of ecological rationality, it is rational to use it in such circumstances. This rationality assessment in terms of accuracy reveals the consequentialist nature of *EBR*. This is where bounded rationality meets epistemic reliabilism.

3. Epistemic reliabilism

In the late 1970s, Alvin Goldman put forth his first reliabilist theory. Since then, he and other authors have kept developing that original proposal, which has been amended by numerous alternative accounts[8]. Initially introduced as a normative theory of justification, it has also been presented as a theory of rationality (Goldman, 2008, p. 240). A common feature of the original 'simple' theory and more sophisticated versions is that the rationality of a belief depends on the reliability of the processes which give rise to the belief in question. In this sense, *PR* is a historical or genetic theory that makes the rationality of a belief depend on the belief's provenance, namely, the process that generates it. This also makes it an indirect form of epistemic consequentialism: the rightness of belief is determined by the reliability of the process which generates the belief and not by its consequences, as maintained instead by direct consequentialism (cf. Ahlstrom-Vij & Dunn, 2018, p. 6).

Epistemic reliabilism rests on process reliability. In the words of Goldman, "a reasoning process is rational if and only if it is *reliable*—i.e., usually generates true beliefs rather than false ones" (Goldman, 2008, p. 240). This property characterises all rational processes and distinguishes them from

[8] For an overview of reliabilist epistemology, see Goldman and Beddor (2021).

irrational ones. More fundamentally, reliabilism assumes that accuracy, or truth-conduciveness, is the primary epistemic goal the reasoner ought to pursue. Beliefs owe their epistemic goodness to the rationality of their belief-forming processes. Thus, simple *PR* consists of the following claim:

> A belief is rational if and only if it is the output of a belief-forming process that is reliable—i.e., one that usually generates true beliefs rather than false ones.

Belief-forming processes confer rationality if they have a high truth ratio, that is, most of the beliefs they produce are true. Precisely how high the truth ratio must be is left vague. The truth-ratio threshold need not be as high as 1, but it must be greater (presumably much greater) than 0.50.

PR distinguishes between justified and unjustified beliefs according to the reliability of the processes that generate them. In his writings, Goldman provides the following examples, we can assume without claiming completeness for them:

> **Unreliable processes**: Confused reasoning, wishful thinking, reliance on emotional attachment, mere hunch or guesswork, hasty generalisation (Goldman, 1979, p. 9), and failure to take account of all one's relevant evidence (or failure to take account of obviously relevant evidence) (Goldman, 1986, p. 104).

> **Reliable processes**: Standard perceptual processes, remembering, good reasoning, introspection (Goldman, 1979, p. 10), and certain patterns of deductive and inductive reasoning (Goldman, 1986, pp. 103-104).

By using these and similar reliable processes and avoiding the use of unreliable ones, the reasoner satisfies the necessary condition for producing rational beliefs. By doing so, reasoners will draw more accurate inferences than they would have done by relying on non-truth-conducive tools.

These are the essential features of simple *PR*. Epistemologists have raised objections to it, and more elaborated versions have been proposed to evade them. To assess the match of *PR* with bounded rationality, we can, for the moment, stick to Goldman's simple version. Later, I will introduce *normality reliabilism*, one of its developments, and show how it copes with some objections to the original proposal.

Now, a tempting inference is that every agent reasoning along the lines prescribed by this theory also fulfils the principle of ecological rationality presented above. Indeed, we should expect reliable processes to perform better than unreliable ones, scoring higher on accuracy. However, the distance between simple *PR* and *EBR* is not as short as one might imagine or hope. Some findings of the psychological sciences regarding the working and effectiveness of inferential methods are hard to square with the prototypical instances of reliable and unreliable processes. In particular, inferential processes that do not consider all available evidence can produce true beliefs contrary to what is assumed by simple reliabilists. This points to a significant normative difference between simple *PR* and *EBR*.

4. Absolute versus contextual assessment of rationality

In the previous sections, I described the main characteristics of *EBR* and *PR*. To sum up: (i) both programmes advance a consequentialist account of epistemic rationality, understood as the achievement of some epistemic values; and (ii) among these values, accuracy, or truth-conduciveness, is the most relevant. At this point, someone might object that accuracy is not all that matters for bounded rationality theorists. This is indeed the case: they also measure the success of cognitive strategies in the world by other criteria, such as speed and frugality. This point requires clarification.

EBR theorists make rationality assessments using a variety of criteria. Looking at these assessments, one notices two recurring points. First, some criteria, such as speed and frugality, are used only occasionally. Moreover, one criterion is never absent in evaluations of rationality: accuracy. Hence, it seems plausible to infer that accuracy is a necessary condition for ecological epistemic rationality, whereas speed and frugality are not. If this is true, we should conclude that accuracy is the most relevant rationality criterion for *EBR*[9].

We can also add summary point (iii): both theories of rationality are based upon the use of appropriate reasoning methods, referred to as 'processes' by reliabilists and 'strategies' by psychologists of bounded rationality. These methods are algorithmic procedures for deriving inferences and making decisions. This is to say, the methods arrive at outcomes by following rule-based patterns of reasoning. This is the case for complex, highly resource-consuming methods and also for fast and frugal heuristics[10].

These are the relevant similarities between *PR* and *EBR*; but we can now ask: Does *PR* perfectly fit *EBR*? In response, it is possible to point out at least two dissimilarities between the two. If we recall the definitions of ecological rationality and process reliability, we will notice the first difference. The former is a comparative notion of rationality: a strategy is ecologically rational compared to some other strategy when the former outperforms the latter on some evaluation criterion. For example, when used in non-compensatory environments, TTB is ecologically rational compared to other inference models, such as WADD. By contrast, process reliability does not require comparison: a reasoning process is

---

[9] This is why a monist epistemology seems most suitable for *EBR*. However, I do not deny that a pluralist epistemology would fit bounded rationality equally well; one might claim, for instance, that accuracy is insufficient and give the other criteria more relevance. This would require explaining whether there is an open or closed list of criteria and the relationships between them, among other things. There is room for future research here.

[10] Heuristics are sometimes seen as mere hunches or anarchic reasoning under the influence of some local stimulus and therefore prone to produce bad outcomes. Although never characterised in these terms, this is partly the image of heuristics that *H&B* has popularised. However, this is not the understanding I adopt here. The example I provided above of the TTB heuristic solving *Creditworthiness* depicts a different image according to which heuristics are step-by-step procedures for computing certain functions, with each step itself being a function. Moreover, heuristics can draw accurate inferences, sometimes even more accurate than those produced by more complex and sophisticated methods.

rational if and only if it is reliable[11]. Therefore, strictly speaking, a reliable process is not ecologically rational unless there is evidence that it outperforms rival processes. But what about the reverse: Is an ecologically rational process also reliable? The definition of ecological rationality does not allow us to say that it is: ecological rationality poses no requirement regarding the truth-conduciveness of a process. However, I think we must say that ecologically rational processes are also truth conducive, otherwise ecological rationality would get into trouble. Let me illustrate this claim with an example.

Imagine two reasoning strategies, *S* and *S'*, whose truth ratios are 0.30 and 0.20, respectively. If ecological rationality were merely a matter of outperforming, then we should conclude that *S* is (ecologically) rational, for *S* produces more true beliefs than *S'*. However, nobody would evaluate *S* as rational, given its low level of accuracy. Thus, there seems to be a hidden condition that ecological rationality requires a minimum level of accuracy, and plausibly this must be much greater than 0.50, as it is for *PR*. Further support comes from the study of simple heuristics. By scrutinising the *EBR* literature, one can see that no heuristic with an average accuracy of less than 0.50 (or probably less than 0.60 or 0.70) has been put forward. When assessed as ecologically rational, heuristics are always remarkably accurate (for a recent systematic review of the literature, see Katsikopoulos et al., 2018).

So, if ecologically rational processes must reach a high truth-ratio threshold, they are reliable. However, the opposite cannot automatically be said to be the case: not all reliable processes are ecologically rational. This discrepancy does not have troublesome consequences for the match between *PR* and *EBR*; and bearing this asymmetry in mind, we can safely defend a strong compatibility between the two programmes.

[11] Notice that there is a sense in which reliability *is* a comparative notion: certain processes produce more correct beliefs than others. For instance, the visual belief that a dog is in front of me formed from detailed and unhurried scanning is plausibly more correct than the visual belief formed from quick and hasty scanning. Accordingly, the former belief is more justified than the latter (Goldman, 1979, p. 10). However, this is not the sense of comparativeness I am using here; my point is that a process can be reliable without this property being established by any comparison with other processes.

There is a second difference between *PR* and *EBR*. Recall the original definition of reliability again: it is the tendency of a process to generate true beliefs rather than false ones. *Where* this tendency needs to be shown or *in which reasoning context*, is a piece of information that the definition of simple reliability does not require. Simple reliabilism is based on a non-located notion of rationality. Now, recall how ecological rationality is defined: a reasoning strategy is ecologically rational *in environment E* compared to some other mechanism when it outperforms this other mechanism on some criterion of comparison. Here, the rationality assessment is relativised to a specific context or environment of reasoning; whereas, as we have just recalled, this relativisation is absent in Goldman's reliabilism.

This discrepancy creates trouble. To see why the non-locatedness of *PR* is a problem for its match with *EBR*, consider once again *Creditworthiness*. When inferring which of the two companies is the more creditworthy, people can either use WADD, which processes all the relevant cues about the two companies, or TTB, which only relies on a fraction of the available data. Evaluating these processes according to simple reliabilism, one should conclude that TTB is irrational. Indeed, as we saw before, not taking into account all one's relevant evidence is considered a paradigmatic example of an *un*reliable process. However, empirical research in cognitive psychology has shown that TTB can predict which company is more creditworthy as accurately as WADD can, and sometimes even more *in non-compensatory environments*, thus proving to be ecologically rational. Therefore, *PR* considers irrational a reasoning process that is considered rational in a specific context by *EBR*, although the two accounts start from a very similar concept of rationality[12].

[12] The empirical evidence gathered by proponents of *EBR* about the accuracy of TTB proves Goldman's assumption that one needs to "take account of all one's relevant evidence" incorrect in situations of uncertainty, where some possible states, their consequences, or their probabilities are not known for sure. Here, less evidence can be more beneficial. However, Goldman's assumption is correct in situations of risk, where all possible states, their consequences, and their probabilities are known. See §6.1 for more on the distinction between uncertainty and risk. I thank an anonymous reviewer for suggesting this caveat.

This dissimilarity highlights how simple *PR* adopts a relatively rigid notion of reliability. According to this, a reasoning process is rational if and only if it tends to produce true beliefs rather than false ones. Apparently, it is tacitly assumed that the process ought to do so wherever it is used. I will call this an *absolute* understanding of process reliability. By not specifying where a process needs to produce true beliefs to be assessed as rational and assuming that this is in every place it is used, simple *PR* precludes the possibility that a method is truth-conducive in specific reasoning contexts and not truth-conducive in others. However, the reliability of a process can change from one context to another, as shown by *Creditworthiness*. Therefore, simple *PR* is not equipped to account for such changes in the performance of reasoning processes from one environment to another. This is a substantive theoretical issue which deserves attention.

Although *PR* does not perfectly fit *EBR*, I believe that something can be done to remedy this; something that is worth doing. Indeed, there are good reasons for preserving their match. On the one hand, *PR* can strengthen the theoretical underpinnings of *EBR*, securing it from puzzling rationality assessments. As I have shown, the purely comparative evaluation *EBR* offers allows for cases of ecologically rational but non-truth-conducive processes. To avoid this pitfall, *EBR* should also require a high truth ratio. This leads to a hybrid definition of ecological rationality according to which a reasoning mechanism M is ecologically rational in environment E in comparison to some other mechanism M' only if M is more accurate than M', and M exceeds a relatively high accuracy threshold.

On the other hand, the psychology research pursued by the ecological project might be seen as a practical implementation of reliabilism. Moreover, *EBR* can also help to improve some conceptual aspects of the epistemic theory proposed by Goldman and other reliabilists, and contribute to overcoming the problem of absolute reliability, as we will see in §6.

So, to preserve the match, simple *PR* should be modified. It seems clear that it needs to acquire the capacity to relativise reliability assessment to the contexts in which the belief-forming processes are used. I will now consider a couple of ways to do this. The first possible solution I contemplate is using a tool the reliabilist family already possesses.

5. Reliability under normal conditions

One of the most disputed questions facing reliabilism concerns the domain in which a process is assessed for reliability. Much of the dispute originated from the so-called 'new evil demon problem' (Cohen, 1984). According to this problem, it is possible to imagine a world where an evil demon creates non-veridical perceptions of physical objects in everybody's mind, which are qualitatively identical to our own, but false in the imagined world. Therefore, the perceptual processes of the inhabitants of this world are unreliable, and their beliefs so caused are unjustified. However, given that their perceptual experiences are qualitatively identical to ours, those very beliefs in the demon world should be justified. Recently, some authors proposed a solution to this problem by taking the domain in which reliability is assessed as the 'normal conditions' for using a given process. Since these conditions are free from evil demons, perceptual and other processes are indeed reliable in these domains.

Normality-based versions of reliabilism as a theory of justified belief can be found, for instance, in Jarrett Leplin (2007, 2009) and Peter J. Graham (2012, 2017). Meanwhile, Beddor and Pavese (2020) introduced a 'normal conditions' variant of the reliabilist account of knowledge. This account offers a simple and effective tool to handle those cases in which knowledge is generated in contexts we would be hard pushed to consider suitable for its production. With some adjustment, this tool can also be used in a theory of rational belief. Moreover, it might offer a solution to the problem of absolute reliability.

Consider the following example by Beddor and Pavese, adapted to a theory of rationality. Temp forms beliefs about the room temperature by consulting a broken thermometer. Unbeknownst to him, Temp has a guardian angel in the room who manipulates the thermostat, ensuring that the room's temperature matches the reading displayed on the thermometer. Any belief Temp forms about the temperature will be true. However, intuitively, Temp's thermal beliefs are not rational.

This case represents an objection to the sufficiency of a reliability condition for rational belief. This objection can be blocked by restricting the domain of reliability assessment to *normal conditions*, understood by Beddor and Pavese as those we would consider 'fair' for performing and assessing the task[13]. So, to defend reliabilism, we might want to modify simple reliability by adding the normality condition:

NORMAL CONDITIONS RELIABILITY

A reasoning process is rational if and only if it is reliable in normal conditions—i.e., it usually generates true beliefs rather than false ones in such conditions.

Accordingly, *Normal Conditions Reliabilism* (*NCR*, from now on) holds that a belief is rational if and only if produced by a process which is reliable in normal conditions. Now, let us consider Temp once more. The task being performed is *forming beliefs about the temperature on the basis of a thermometer*. The reasoning process used is consulting a broken thermometer. Intuitively, having a hidden helper manipulating the thermostat is a highly abnormal circumstance for Temp's task. However, in normal conditions, namely, in helper-free worlds, consulting a broken thermometer leads to false beliefs. So, Temp's reasoning process is unreliable in normal conditions. Therefore, *NCR* deems Temp's beliefs to be irrational.

Unlike Temp's case, the new evil demon problem challenges the necessity of the reliability condition for rational belief. Can *NCR* block this criticism as well? Consider Danny, a subject deceived by the evil demon, who possesses false perceptions of physical objects identical to ours (Danny could alternatively be a subject in a vat whose brain is stimulated by a supercomputer which generates non-veridical representations of the world). The task Danny is performing is *forming beliefs on the basis of visual stimuli* or simply *forming visual beliefs*. To do so, the perceptions of physical objects are relied upon. Intuitively, the normal conditions for Danny's task are free from the demon's interventions. Thus, Danny is in highly abnormal conditions for performing and assessing the task. In the demon world,

[13] Those authors note that it would be difficult to give a precise and non-circular analysis of what these conditions consist of. However, intuition about cases might suffice here. Indeed, they argue, we often manifest our tacit conception of fair conditions when we evaluate reasoning performance.

Danny's reasoning process is unreliable, and the resultant beliefs are false. In contrast, Danny's reasoning process is reliable in normal conditions. Hence, *NCR* deems Danny's beliefs in the demon world to be rational (as also for the brain-in-vat case).

Normality-based reliabilism is a variant of the traditional theory that Goldman has welcomed (Goldman & Beddor, 2021, pp. 10-11). Its novelty consists in relativising the assessment of process reliability to normal conditions, understood as those we would consider fair for performing and assessing a given task. Now, the questions for us are: Can normality reliabilism account for the change in performance of a reasoning method from one environment to another? Can it explain why using TTB in non-compensatory environments is rational but it is irrational in compensatory ones? And: Does *NCR* succeed where simple reliabilism fails?

Empirical and analytical research has shown that the performance of many reasoning strategies is influenced by features of the reasoning context and the cognitive setup of the reasoner. Non-compensatoriness, for instance, makes TTB as accurate as WADD and sometimes even more accurate. Should we conclude that non-compensatory environments are normal conditions for TTB? These conditions are those we consider fair for performing and assessing a reasoning task. For instance, under these conditions, no manipulation has occurred to make an unreliable process truth-conducive (as occurred in the Temp case) or to make a reliable process non-truth-conducive (as in the evil demon case). However, the normality condition does not seem to solve the TTB case. Indeed, the change in performance of TTB from one environment to another is not due to any manipulation of the reliability of the process. Instead, we would need to say that the performance of TTB is rational in an environment having a specific feature, namely, non-compensatoriness. Fairness for performance and assessment seems to be irrelevant here. Hence, *NCR* fails to offer a solution to the problem of absolute reliability.

Notice that this failure does not mandate abandoning the normality condition, which blocks the new evil demon and the guardian angel objections to simple *PR*. We should not renounce this achievement, given that we propose *PR* as an epistemology for bounded rationality. Therefore, we should embed the normality condition in our definition of process reliability and continue the search for a way to complete the account and solve the problem of absolute reliability.

6. A contextual approach to reliability

In light of the foregoing considerations, relativising reliability assessment is the way to go. Here, in addition to normality reliabilism, inspiration can be found in a recent dispute in the reliabilist camp. The so-called 'temporality problem' for reliabilism (Frise, 2018; Tolly, 2019) consists of the fact that a process can be more reliable at one time than another. For instance, weather forecasting has improved over time. Thus, the process type 'forming a belief based on the forecast' is more reliable now than it was twenty years ago. This raises a question about the temporal parameters we should use when evaluating reliability. Imagine we are evaluating whether belief *B* is justified at time *t*. Should we focus on whether the belief-forming process responsible for *B* is reliable at *t*? Or should we consider whether it has always been reliable up until *t*; or at all times that are temporally close to *t*? Or something else? According to Frise (2018), the temporality problem generates insurmountable difficulties for reliabilism, whereas Tolly (2019) rejects this conclusion and shows that there are reasonable temporal parameters for the reliabilist to adopt.

I do not need to enter into the details of the issue. Nevertheless, it provides helpful insight when framing a solution to the problem of absolute reliability. First, it seems to support the idea that an absolute conception of reliability is implausible: it may be the case that most reasoning processes are more or less reliable depending on some contextual features, such as the time or the place of their use. Hence, it encourages the relativisation of reliability. Moreover, the issue suggests that, by analogy with the temporality problem, reliabilism could also have a *spatiality problem*: a process can be more reliable in one environment than another. Thus, a possible way to tackle the question of absolute reliability might be to introduce a *spatial parameter E* which relativises the reliability assessment to specific environments. Reliabilism could borrow this parameter from ecological rationality, which already possesses one. Hence, as a first approximation, an environment is generated by the interplay between features of the reasoning scenario (for instance, available information, time pressure, etc.) and features of the cognitive setup of the reasoner (for example, memory states, computational faculties, etc.).

I propose that the assessment of process reliability should be a function of the process truth ratio, the conditions for performing and assessing the reasoning task (i.e., whether they are normal or abnormal), and the environment in which the process is used. Therefore, the following principle should replace simple reliability:

> NORMAL ENVIRONMENT RELIABILITY
>
> A reasoning process is rational in environment *E* if and only if it is reliable in environment *E* under normal conditions—i.e., it usually generates true beliefs rather than false ones in this environment under normal conditions.

To illustrate this principle, recall the task of assessing the creditworthiness of two companies. The cognitive goal is to make the most accurate judgement based on the available cues. Two reasoning strategies are available to achieve this: TTB and WADD. Cognitive psychology has shown that the former performs as well as the latter and sometimes even better in environments characterised by non-compensatory information. In contrast, TTB performs poorly and is usually outperformed by WADD in environments characterised by compensatory information. The normative conclusion is that TTB, an inferential process that ignores part of the available evidence, is rational in non-compensatory environments since it is reliable there—i.e., it tends to produce true beliefs rather than false ones in these environments. In contrast, it is irrational in compensatory environments since it is unreliable there—i.e., it tends to produce false beliefs rather than true ones in these environments. Thus, normal environment reliability can account for the change in the performance of this reasoning process from some environments to others.

With this new notion of reliability, we can now introduce the corresponding epistemic theory called *Normal Environments Reliabilism* (*NER*, henceforth), according to which a belief is rational in environment *E* if and only if it is the output of a belief-forming process that is reliable in environment *E* under normal conditions—i.e., it usually generates true beliefs rather than false ones in this environment under normal conditions. This is a candidate for an epistemology for *EBR*. Notice that *EBR* only offers a theory of rational *reasoning strategy*: an account of what makes a belief-forming process rational. In this sense, ecological rationality parallels process reliability. However, *EBR* provides no fully fledged epistemology intended as a theory of rational *belief*. A further step is needed, and reliabilism seems helpful here. The beliefs an agent produces under environmental and cognitive constraints are rational in the context of production if and only if they are generated by belief-forming processes that, in normal conditions, are reliable in that context.

Let us consider the main features of the new account. First, it should be clear that normal environment reliability is not more demanding than absolute reliability, as it does not introduce any further rationality criterion. Accuracy, or truth-conduciveness, remains the only standard for evaluating the performance of a reasoning process.

Second, normal environment reliability is a located notion of rationality: it refers to the tendency of a reasoning process to be truth-conducive in specific contexts. Thus, reference must be made to environments when assessing this kind of reliability. To do this, the assessor should use knowledge about where a strategy performs well and where it does not.

Moreover, the parameter *E* and the normality condition play different functions, and we need both. *E* individuates reasoning contexts with features that play out in favour of or against the performance of a reasoning process, such as non-compensatoriness and compensatoriness for TTB. Meanwhile, normality discriminates between normal and abnormal conditions for the performance and assessment of a task. Conditions are normal when nothing has altered the reliability of a process that can be applied to tackle the task. In *Creditworthiness*, conditions are normal if evil-demon-like factors have not modified the reliability of TTB and WADD. Imagine instead that an agent is assisted by her guardian angel who, whenever and wherever she uses TTB, ensures that her prediction matches the actual relative creditworthiness of the two companies. Suppose also that the agent is in a compensatory environment. The reasoner will draw many true inferences thanks to the aid of the guardian angel. Will they be rational? Under normal, guardian-angel-free conditions, TTB is unreliable in compensatory environments. Hence, the agent’s inferences produced by TTB will not be rational. Suppose now that the agent is instead in a non-compensatory environment. In this case, the agent's inferences produced by TTB will be rational because TTB is reliable in non-compensatory environments.

Finally, *NER* is more descriptively plausible than the simple, absolute version. Indeed, it allows us to distinguish where a process is reliable and where it is not, and it does not oblige us to consider a process reliable or unreliable wherever it is used. This brings justice to the empirical evidence gathered by *EBR*.

Reliability can now be relativised to environments where reasoning methods are used or usable, making the environment a central notion in the new epistemic framework. So, at this point, it might be opportune to clarify how this notion should be understood.

6.1. Reasoning environments

Reasoning environments are contexts within which and upon which the reasoner acts. Their features affect the agent's thoughts and actions in various ways. To begin with, they provide the input processed by reasoners and shape the tools available for reaching their goals. Moreover, reasoning environments influence process reliability. As we saw in *Creditworthiness*, the reliability of processes such as TTB and WADD sensibly varies depending on the environment in which the reasoning takes place. Therefore, the reliability assessment cannot neglect environmental factors.

Offering a complete characterisation of reasoning environments might be difficult, but it is possible to highlight several important structures. Gigerenzer and Sturm (2012) isolated three: the degree of uncertainty, the number of alternatives, and the learning sample size. This list is not meant to be complete; other structural features might be added. Time pressure is a candidate here. For instance, Marewski and Schooler (2011) show that the inferential accuracy of heuristics varies depending on whether people can take all the time they want to draw the inference or are given a short timeframe. But I will now briefly address the three items already on the list.

Uncertainty. The degree of uncertainty refers to the extent to which available cues can predict a criterion. Some predictions are more uncertain than others, depending on what is predicted. For instance, the future performance of stocks and funds is highly unpredictable; heart attacks are slightly more predictable; and tomorrow's weather is even more predictable. Uncertainty characterises most reasoning environments in which humans can find themselves. According to the definition introduced by economist Frank Knight (1921), it applies to all those situations where we cannot know all the information about a future event and, therefore, cannot calculate its probability accurately. It must be distinguished from *risk*, which is the feature of those situations where, although we do not know the outcome of a future event, we can accurately measure its probability. A pure game of chance in a casino is an example of a risky situation.

Number of alternatives. When drawing inferences or making decisions, people can deal with different numbers of alternatives. These can refer to individual objects (such as companies) or sequences (such as moves or pathways). This environmental structure becomes particularly relevant when the alternatives are many. Algorithmic procedures for estimating a criterion might run into the problem of computational intractability, the impossibility of estimating the criterion optimally, given a massive number of alternatives. Consider, for instance, the game of chess. Although an optimal, best sequence of moves does, in theory, exist, no mind or computer can determine it. This is one of the reasons why both machines and humans rely on simpler, non-optimisation techniques, including heuristics.

Sample size. This structural feature of the environment constitutes the number of sampling units included in the sample. Statistical methods for predicting future events estimate parameter values from past data. Sample size directly influences the performance of predictive strategies. For example, consider two models, one complex with many free parameters and a heuristic with few parameters, predicting an uncertain future event using a small sample of data. In this circumstance, the resulting error due to 'variance' committed by the complex model may exceed the error due to the heuristic's 'bias', making the latter strategy preferable (Gigerenzer & Brighton, 2009). Thus, different sample sizes can call for different predictive methods.

Notice that 'environment' is not used here to indicate only the physical environment. For instance, the degree of uncertainty reflects the environment as well as the limited understanding of the mind. Hence, the degree of uncertainty is located in the mind–environment system (Gigerenzer & Sturm, 2012, pp. 256-258). More generally, the environment here is to be understood as the totality of the relevant physical features of the world that, combined with those of the reasoner's cognitive system, favour the performance of some reasoning strategies over that of others.

7. Conclusion

I have explored the relationship between reliabilist epistemology and the ecological approach to bounded rationality. The two accounts have been recognised as largely compatible by the authors who have contributed most to their development. However, I first show that the comparative definition of ecological rationality allows an unreliable process to be assessed as rational. To avoid this problem, I suggest that *EBR* might adopt a reliability condition. Furthermore, I highlight how the absolute character of Goldman's simple reliabilism does not square with the located notion of rationality adopted by *EBR*. Simple *PR* cannot account for processes that are reliable in one reasoning context and unreliable in others, the existence of which is an empirical fact widely reported in the psychology literature.

A first attempt to save the relationship between the two theories resorted to normality reliabilism, which relativises process reliability to normal conditions, understood as 'fair' domains for performing and assessing a reasoning task. Although the normality condition solves the new evil demon problem for reliabilism, it does not capture those environmental features that play out in favour of or against a reasoning strategy. Thus, I proposed solving the problem of absolute reliability by relativising process reliability to reasoning environments. Normal environment reliability adds the normality condition and the environmental parameter *E* to simple reliability. The former blocks some standard objections to simple *PR*[14]; the latter individuates reasoning environments whose features influence process reliability.

---

[14] The normality condition might not block all objections to simple reliabilism. Consider Bonjour's (1980) clairvoyance problem: Norman, a completely reliable clairvoyant with no evidence of his clairvoyant powers, does not seem to be justified in his beliefs. However, the conditions for performing and assessing his reasoning task seem normal, so *NER* must conclude that Norman's beliefs are rational. Several solutions to the clairvoyance

A new theory of rational beliefs emerges, *NER*, which, as opposed to the original reliabilist view, does justice to those reasoning processes that are truth-conducive in some environments and non-truth-conducive in others, such as TTB in non-compensatory and compensatory environments, respectively.

We can see an advantage of *NER* when it comes to heuristic reasoning. As we saw at the beginning, there is disagreement about the pros and cons of heuristics. We can interpret the ongoing debate as a controversy about the *reliability* of heuristics, since scholars disagree about the tendency of heuristics to produce true beliefs rather than false ones. For instance, Goldman expressed his sympathy for the line of thought according to which heuristics are "highly *error-prone*, or indeed *biased toward error*", the "quick and dirty" view often associated with Tversky and Kahneman's *H&B* approach (Goldman, 2017, p. 24). However, this view clashes with the empirical findings of Gigerenzer and his colleagues of the *EBR*, which have instead shown when heuristics perform well and when they do not. Adopting a contextual approach to epistemic rationality might contribute to solving the dispute about heuristics since relativising rationality assessment to reasoning environments leads to more balanced conclusions. Indeed, using normal environment reliability, it is possible to express that a particular heuristic is reliable in some environments and unreliable in others. Accordingly, agents are rational if they use such a heuristic in the former environments and irrational if they do so in the latter.

On the one hand, I hope to have contributed to strengthening the philosophical underpinnings of *EBR* by exploring and critically assessing its relationship with *PR*. On the other hand, I believe that reliabilism can now account for empirical facts that are well-known in cognitive sciences. If this is indeed the case, this work can represent a starting point for future research across philosophy and psychology of reasoning.

---

problem have been proposed, including combining reliabilism with evidentialist elements, appealing to primal systems, and attribution theories (see Goldman & Beddor [2021] for an overview of these solutions). Here I will neither choose one of these solutions nor offer my own: defending reliabilism from all criticism goes beyond the goal of this paper, which is to sketch a reliabilist account of bounded rationality. However, future research might strengthen the current account.

Acknowledgements

I am grateful to Konstantinos Katsikopoulos, Thomas Sturm, and David Thorstad for their comments on earlier drafts of this paper. Special thanks to Nick Hughes for his comments and crucial discussions. Thanks also to Michele Palmira and Sven Rosenkranz for helpful conversations and to the audiences of the Logos Epistemology Reading Group, the Logos Graduate Reading Group, and the Adaptive Behavior and Cognition Workshop for incisive questions. Thanks to Christopher Evans for proofreading the final version.

Statements and declarations

This work was supported by the Secretariat for Universities and Research of the Catalonian Department of Business and Knowledge and the European Social Fund.

The author has no relevant financial or non-financial interests to disclose.

References

Ahlstrom-Vij, K., & Dunn, J. (Eds.). (2018). *Epistemic consequentialism.* Oxford: Oxford University Press.

Aikman, D., Galesic, M., Gigerenzer, G., Kapadia, S., Katsikopoulos, K., Kothiyal, A., . . . Neumann, T. (2014). *Taking uncertainty seriously - simplicity versus complexity in financial regulation.* Financial Stability Paper No. 28. London, UK: Bank of England.

Beddor, B., & Pavese, C. (2020). Modal Virtue Epistemology. *Philos Phenomenol Res, 101*, 61-79. doi:https://doi.org/10.1111/phpr.12562

Bishop, M., & Trout, J. (2005). *Epistemology and the Psychology of Human Judgment.* New York: Oxford University Press.

BonJour, L. (1980). Externalist Theories of Empirical Knowledge. *Midwest Studies In Philosophy, 5*(1), 53-74. doi:https://doi.org/10.1111/j.1475-4975.1980.tb00396.x

Brighton, H., & Gigerenzer, G. (2015). The bias bias. *Journal of Business Research, 68*, 1772–1784. doi:http://dx.doi.org/10.1016/j.jbusres.2015.01.061

Brighton, H., & Todd, P. M. (2009). Situating rationality: Ecologically rational decision making with simple heuristics. In P. Robbins, & M. Aydede, *The Cambridge handbook of situated cognition* (p. 322–346). Cambridge University Press.

Cohen, S. (1984). Justification and Truth. *Philosophical Studies, 46*(3), 279-295. doi:https://doi.org/10.1007/BF00372907

Feldman, R. (2012). *Naturalized Epistemology.* (E. Zalta, Ed.) Retrieved December 9, 2022, from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/sum2012/entries/epistemology-naturalized/>

Frise, M. (2018). The reliability problem for reliabilism. *Philosophical Studies*, *175*, 923–945. doi:https://doi.org/10.1007/s11098-017-0899-0

Gigerenzer, G. (1991). How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases". *European Review of Social Psychology, 2*(1), 83-115. doi:https://doi.org/10.1080/14792779143000033

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, *103*(3), 592–596. doi:https://doi.org/10.1037/0033-295X.103.3.592

Gigerenzer, G. (2008). *Rationality for Mortals. How People Cope with Uncertainty.* Oxford University Press.

Gigerenzer, G. (2019). Axiomatic rationality and ecological rationality. *Synthese*, *198*, 3547–3564. doi:https://doi.org/10.1007/s11229-019-02296-5

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science, 1*(1), 107-143. doi:https://doi.org/10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., & Goldstein, D. (1996). Reasoning the Fast and Frugal Way: Models of Bounded Rationality. *Psychological Review, 103*(4), 650-669. doi:https://doi.org/10.1037/0033-295x.103.4.650

Gigerenzer, G., & Sturm, T. (2012). How (far) can rationality be naturalized? *Synthese, 187*, 243–268. doi:10.1007/s11229-011-0030-6

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The Foundations of Adaptive Behavior.* Oxford University Press. doi:https://doi.org/10.1093/acprof:oso/9780199744282.001.0001

Gigerenzer, G., Todd, P., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart.* Oxford University Press.

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

Goldman, A. (1979). What is Justified Belief? In G. S. Pappas, *Justification and Knowledge* (Vol. 17, pp. 1-23). Dordrecht: Springer. doi:https://doi.org/10.1007/978-94-009-9493-5_1

Goldman, A. (1986). *Epistemology and Cognition.* Harvard University Press.

Goldman, A. (1999). *Knowledge in a social world.* Oxford: Oxford University Press.

Goldman, A. (2008). Human Rationality: Epistemological and Psychological Perspectives. *Philosophie: Grundlagen und Anwendungen*, 259-276. doi:https://doi.org/10.30965/9783969750056_017

Goldman, A. (2017). What Can Psychology Do for Epistemology? Revisiting Epistemology and Cognition. *Philosophical Topics, 45*(1), 17–32. doi:https://www.jstor.org/stable/26529423

Goldman, A., & Beddor, B. (2021, May 21). *Reliabilist Epistemology*. (E. N. Zalta, Editor) Retrieved December 16, 2022, from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/sum2021/entries/reliabilism/>

Graham, P. (2012). Epistemic Entitlement. *Noûs, 46*(3), 449-482. doi:https://doi-org.are.uab.cat/10.1111/j.1468-0068.2010.00815.x

Graham, P. (2017). Normal Circumstances Reliabilism: Goldman on Reliability and Justified Belief. *Philosophical Topics*, *45*(1), 33-62. doi:https://www.jstor.org/stable/26529424

Hertwig, R., Leuker, C., Pachur, T., Spiliopoulos, L., & Pleskac, T. (2021). Studies in Ecological Rationality. *Topics in Cognitive Science*, 1–24. doi:10.1111/tops.12567

Jenny, M., Pachur, T., Williams, S., Becker, E., & Margraf, J. (2013). Simple rules for detecting depression. *Journal of Applied Research in Memory and Cognition, 2*(3), 149–157. doi:https://doi.org/10.1037/h0101797

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237–251. doi:https://doi.org/10.1037/h0034747

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*(2), 582-91. doi:10.1037/0033-295x.103.3.582

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

Katsikopoulos, K. (2014). Bounded rationality: the two cultures. *Journal of Economic Methodology, 21*(4), 361-374. doi:https://doi.org/10.1080/1350178X.2014.965908

Katsikopoulos, K. V., Durbach, I. N., & Stewart, T. J. (2018). When should we use simple decision models? A synthesis of various research strands. *Omega*, *81*, 17-25. doi:https://doi.org/10.1016/j.omega.2017.09.005

Kitcher, P. (1992). The Naturalists Return. *The Philosophical Review, 101*(1), 53-114. doi:https://doi.org/10.2307/2185044

Knight, F. (1921). *Risk, uncertainty and profit.* Boston: Houghton Mifflin Co.

Kornblith, H. (2014). *A Naturalistic Epistemology: Selected Papers.* Oxford: Oxford University Press.

Leplin, J. (2007). In defense of reliabilism. *Philosophical Studies*, *134*, 31–42. doi:10.1007/s11098-006-9018-3

Leplin, J. (2009). *A Theory of Epistemic Justification.* Netherlands: Springer Dordrecht. doi:https://doi.org/10.1007/978-1-4020-9567-2

Marewski, J. N., & Schooler, L. J. (2011). Cognitive Niches: An Ecological Model of Strategy Selection. *Psychological Review, 118*(3), 393–437. doi:10.1037/a0024143

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision, 52*(1), 29–71. doi:https://doi.org/10.1023/A:1015516217425

Piattelli-Palmarini, M. (1994). *Inevitable illusions: how mistakes of reason rule our minds.* New York: Wiley.

Samuels, R., Stich, S., & Bishop, M. (2002). Ending the Rationality Wars How to Make Disputes about Human Rationality Disappear. In R. Elio (Ed.), *Common Sense, Reasoning, and Rationality* (pp. 236–268). New York. doi:https://doi.org/10.1093/0195147669.003.0011

Schurz, G., & Hertwig, R. (2019). Cognitive Success: A Consequentialist Account of Rationality in Cognition. *Topics in Cognitive Science, 11*(1), 7-36. doi:https://doi.org/10.1111/tops.12410

Serwe, S., & Frings, C. (2006). Who will Win Wimbledon? The Recognition Heuristic in Predicting Sports Events. *Journal of Behavioral Decision Making, 19*(4), 321–332. doi:https://doi.org/10.1002/bdm.530

Simon, H. (1957). *Models of Man.* New York: John Wiley.

Simon, H. (1976). From substantive to procedural rationality. In T. J. Kastelein, S. K. Kuipers, W. A. Nijenhuis, & G. R. Wagenaar, *25 Years of Economic Theory.* Boston, MA: Springer. doi:https://doi.org/10.1007/978-1-4613-4367-7_6

Simon, H. (1990). Invariants of Human Behavior. *Annual Review of Psychology, 21*, 1-20. doi:https://doi.org/10.1146/annurev.ps.41.020190.000245

Şimşek, Ö. (2013). Linear decision rule as aspiration for simple decision heuristics. *Advances in neural information processing systems, 26*, 2904–2912.

Stein, E. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science.* Oxford: Clarendon.

Sturm, T. (2012). The “Rationality Wars” in Psychology: Where They Are and Where They Could Go. *Inquiry, 55*(1), 66-81. doi:https://doi.org/10.1080/0020174X.2012.643628

Sturm, T. (2019). Formal versus Bounded Norms in the Psychology of Rationality: Toward a Multilevel Analysis of Their Relationship. *Philosophy of the Social Sciences, 49*(3), 190–209. doi:https://doi.org/10.1177/0048393119842786

Sturm, T. (2020). Philosophical naturalism and bounded rationality. In R. Viale (Ed.), *Routledge Handbook of Bounded Rationality* (pp. 73-89). Routledge.

Todd, P., & Gigerenzer, G. (2012). *Ecological Rationality: Intelligence in the World.* Oxford University Press.

Tolly, J. (2019). Does reliabilism have a temporality problem? *Philosophical Studies, 176*, 2203–2220. doi:https://doi.org/10.1007/s11098-018-1122-7

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293-315. doi:https://doi.org/10.1037/0033-295X.90.4.293

Vranas, P. (2000). Gigerenzer's normative critique of Kahneman and Tversky. *Cognition, 76*(3), 179-93. doi:10.1016/s0010-0277(99)00084-0

Wheeler, G. (2020, July 28). *Bounded Rationality.* (N. Zalta, Ed.) Retrieved July 27, 2023, from The Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/archives/fall2020/entries/bounded-rationality/>