

How (far) can rationality be naturalized?

Gerd Gigerenzer · Thomas Sturm

Received: 30 September 2011 / Accepted: 30 September 2011 / Published online: 1 November 2011
© Springer Science+Business Media B.V. 2011

Abstract The paper shows why and how an empirical study of fast-and-frugal heuristics can provide norms of good reasoning, and thus how (and how far) rationality can be naturalized. We explain the heuristics that humans often rely on in solving problems, for example, choosing investment strategies or apartments, placing bets in sports, or making library searches. We then show that heuristics can lead to judgments that are as accurate as or even more accurate than strategies that use more information and computation, including optimization methods. A standard way to defend the use of heuristics is by reference to accuracy-effort trade-offs. We take a different route, emphasizing ecological rationality (the relationship between cognitive heuristics and environment), and argue that in uncertain environments, more information and computation are not always better (the “less-can-be-more” doctrine). The resulting naturalism about rationality is thus normative because it not only describes what heuristics people use, but also in which specific environments one should rely on a heuristic in order to make better inferences. While we desist from claiming that the scope of ecological rationality is unlimited, we think it is of wide practical use.

Keywords Rationality · Judgment and decision-making · Heuristics · Uncertainty

It is often doubted that rationality can be naturalized. Is it not preposterous to derive norms of reasoning from descriptive statements of empirical science? How is it possible to explain norms scientifically? And is reason not in part the source of a priori, strictly universal knowledge claims? The first of these three skeptical questions

G. Gigerenzer (✉)
Max Planck Institute for Human Development, Berlin, Germany
e-mail: gigerenzer@mpib-berlin.mpg.de

T. Sturm
Department of Philosophy, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain
e-mail: tsturm@mpiwg-berlin.mpg.de

relies on a distinction, prominently advocated by Kant, Frege, the logical empiricists, Popper, and many others: How human beings actually reason is one issue, to be studied empirically; how they ought to reason is another, inappropriate for empirical research. This is–ought distinction is reflected in a division of labor between psychology and philosophy. In this view, norms of reasoning cannot be derived from empirical psychology. The second question concerns the explanation of norms of reasoning by empirical science, for example, by evolutionary theory (Chiappe and Vervaeke 1997). The third question addresses a version of naturalism that rejects the distinction between a priori and empirical knowledge claims or asserts, for instance, that basic principles of logic may be empirically revised (Putnam 1982).

In this essay, we do not directly respond to any of these doubts. Instead, we outline how a naturalization of rationality is in fact possible, rather than merely promised from the philosophical armchair. We will rely on recent psychological research on human reasoning. We describe theoretical, conceptual, and methodological assumptions involved in this research. We also make clear *how far* this naturalization of rationality goes, in order to abstain from any universal claims. Two points may be made in this respect.

First, in our aim to naturalize rationality, we do not claim that the normative inquiries of epistemology or decision theory should be replaced by purely descriptive inquiries into actual human reasoning and decision making. Such a move was advocated by Quine (1969) in his original call for a naturalized epistemology (for his more nuanced later claims, see Quine 1990, p. 20). It has rightly been objected that this proposal amounts to changing the topic (Kim 1988). Our naturalism, however, is not a “replacement” but a normative naturalism. Second, while we intend to show that norms of rationality can be derived from empirical research, we do not assert that the gap between empirical and normative inquiries can be bridged always and everywhere. In the debate about naturalism in epistemology, for instance, some defend a radically universal normative naturalism: *All* questions of standard epistemology about how to justify or revise our knowledge claims should either be solved by using the empirical sciences, especially cognitive psychology, or else be abandoned (Bishop and Trout 2005; Bishop 2008). In contrast, our naturalism is distinctive in that it sets itself limits: In some important domains, one can infer from empirical research what norms of rationality are best, as well as how human reasoning can be improved. In other domains one cannot; that is, in these the “standard” conception of rationality (Stein 1996) as being based upon certain rules of logic or probability is not undermined by our arguments.

One background reason for this limitation is that rationality can be defined in several ways. To illustrate, if one claims that an X (a rule, a person, or a person’s capacities or behaviors) is rational, this may exclude either that X is nonrational or that X is irrational. Aristotle famously defined man as a rational animal, thereby distinguishing human beings from other animals. However, he did not mean that cats or dogs are irrational. Rather, they—or their behavior—are neither rational nor irrational. Here are two basically different senses of ‘rationality,’ one opposed to *arationality* and the other opposed to *irrationality*. Moreover, and more importantly, the specific standards according to which one can describe a person as being rational in the sense opposed to irrationality are disputed. There are also further distinctions one might draw: between

theoretical and practical rationality, between means-ends (or instrumental) and principled rationality, between procedural and substantive rationality, between the rationality of rules or the rationality of persons, and so on (Mele and Rawlings 2004; Hanna 2006), although these are largely irrelevant here (but see Sect. 5).

The rationality naturalized by our psychological research program picks out one specific sense of rationality, albeit an important one. We start from a kind of means-ends rationality—the “default” notion of rationality (Nozick 1993, p. 133): to know by which means to achieve one’s ends always requires empirical knowledge. But our naturalism about rationality is not merely based upon this triviality. It takes a distinctive and substantive further step by claiming that (1) humans rely on heuristics to solve problems, and (2) heuristics can lead to *more* accurate judgments than strategies that use more information and computation, including optimization methods, if one takes into account the relation between a reasoner’s heuristics and his or her environment. We claim that normative theories of rationality have paid too little, if any, attention to this important relation. Thus, we supplement the instrumental notion of rationality by an “ecological” dimension. The crucial point to which this leads is that the claim (2), which has strong empirical evidence in its favor, is *not* merely descriptively relevant. In an uncertain environment, more information and computation are not always better, as we shall argue. This doctrine of “less-is-more”—or more precisely, “less-can-be-more”—is the core of the current attempt to naturalize a particular domain of normative theories of reasoning and decision making. In short, our naturalism about rationality is normative, ecological, and limited. However, it would be a misunderstanding to think that its limitation makes it uninteresting. The cases we shall present of deriving norms from empirical studies of human reasoning will be striking.

We start with an example concerning investments (Sect. 1) before introducing the crucial psychological notion of heuristics and explaining how heuristics should be studied empirically (Sect. 2). Then, we examine the relation between heuristics and the environments in which we reason. Here we shall also determine the limits of the present naturalism by sketching central features of the kinds of environments in which heuristics succeed over standard optimization norms such as the ones embodied in probability or decision theory (Sect. 3). Furthermore, we explain the basic reasons for why and when heuristics are to be preferred over optimization strategies that require more information and calculation. We distinguish two quite different reasons for why one should rely on heuristics (Sect. 4). Finally, we present and reject an objection that might easily arise, namely that we are confusing issues of the justification and application of norms, or that our normative naturalism is built upon ambiguities in the concept of rationality (Sect. 5).

1 Satisficing can lead to better outcomes than optimizing

In 1990, Harry Markowitz received the Nobel Prize in Economics for his path-breaking work on optimal asset allocation. He addressed a vital investment problem faced by anyone saving for retirement or attempting to earn money on the stock market: How should money be invested in N assets? To avoid the risk of putting everything in one basket, one should diversify, but the question of how remains. Markowitz derived

the optimal rule for allocating wealth across assets, known as the *mean-variance portfolio*, because it maximizes the return (mean) and minimizes the risk (variance). Yet when he made his own retirement investments, instead of using his award-winning optimization technique, he relied instead on a simple heuristic, the *1/N rule*:

Invest equally in each of the N alternatives.

Markowitz was not the only one to use this heuristic; empirical studies indicate that about 50% of ordinary people intuitively rely on it (Huberman and Jiang 2006). Skeptics may find this rule naive, however, and ask whether optimization is not the better alternative. To answer such questions, a study compared the *1/N* rule with the mean-variance portfolio and 13 other “optimal” asset allocation policies in seven investment problems, such as allocating one’s money to 10 American industry funds (DeMiguel et al. 2009). The optimization models included sophisticated Bayesian and non-Bayesian models, which were given 10 years of stock data to estimate their parameters and on that basis had to predict the next month’s performance; after this, the 10-year window was moved a month ahead and the next month had to be predicted, and so on, until no data remained. *1/N*, in contrast, ignores all past information. The performances of all strategies were evaluated by three standard financial measures: *1/N* ranked first on certainty equivalent returns, second on turnover, and fifth on the Sharpe ratio, respectively. Despite complex estimations and computations, none of the optimization methods could consistently earn better returns than the simple heuristic.

How can a simple heuristic outperform optimization strategies? Note that in an ideal world where the mean-variance portfolio could estimate its parameters perfectly, that is, without error, it would do best. But in an uncertain world, even with 10 years of data (which is more than the typical investment firm relies on), optimization no longer necessarily leads to an optimal outcome. In an uncertain world, ignoring information leads to better decisions. Yet our point is neither that simple heuristics, or satisficing (Simon 1990), are always better than optimization methods nor that optimization methods are always better, as is typically assumed. No heuristic or optimization strategy is superior in all worlds. The question instead is in what environment is a given heuristic more accurate than an optimization strategy, and when is the opposite true? This is the question of the *ecological rationality* of a heuristic (Gigerenzer and Selten 2001). The answer requires analyzing structures of the environment and the match between heuristic and environment. For the choice between *1/N* and the mean-variance portfolio, three relevant environmental features are:

- (1) degree of uncertainty,
- (2) number N of alternatives, and
- (3) size of the learning sample.

It is difficult to predict the future performance of funds, because uncertainty is high. The size of the learning sample is the estimation window, with 5–10 years corresponding to the typical calibration of portfolio models in investment practice. *1/N* tends to outperform the mean-variance portfolio if uncertainty is high, the number of alternatives is large, and the learning sample is relatively small. This qualitative insight leads to a quantitative question: If there are 50 alternatives, how large a learning sample is needed for the mean-variance portfolio to eventually outperform the simple heuris-

tic? The answer is: 500 years of stock data, assuming that in the year 2500 the same funds—and the stock market—still exist (DeMiguel et al. 2009).

This analysis allows the following statement: The intuitive $1/N$ heuristic is normatively preferable to the mean-variance optimization method (in the sense that it leads to better financial outcomes) for $N = 50$ if there is less than 500 years of stock data. What is normatively required is no longer defined by consistency principles, optimization, or armchair thinking but is based on an empirical analysis. It is also not a statement about what is best, only what is better than alternatives.

2 What heuristics are

As $1/N$ illustrates, a heuristic is a strategy that ignores some information and minimizes computation, which is exactly what is needed in this particular example for better (and faster) judgments. The repertoire of heuristics, their building blocks, and the evolved capacities they exploit can be described as the mind's *adaptive toolbox* (Gigerenzer 2008).

In *Simple heuristics that make us smart* (Gigerenzer et al. 1999), a research program was presented that connects the empirical and the normative studies of human rationality. It has three research goals:

- (1) *The heuristics in the adaptive toolbox* What are the heuristics, their building blocks, and the evolved capacities in the adaptive toolbox of an individual or a species?
- (2) *The connection between heuristics and environment* What environmental structures can a heuristic exploit?
- (3) *Improving reasoning and deciding* How can the answers to (1) and (2) be utilized to design environments and create expert systems that improve judgment and decision making?

Since the appearance of that volume, an increasing number of researchers have contributed to answering these questions, provided evidence for situations where it is indeed rational that people rely on heuristics, and studied the power of heuristics in domains including business, medical diagnosis, and the law (see Gigerenzer 2007, 2008). Let us begin with Question (1) in this section; Questions (2) and (3) will be addressed mainly in Sects. 3 and 4. In Sect. 4, we will also explain more generally when and why heuristics are to be preferred to optimization strategies that require more information and calculation.

2.1 The heuristics in the adaptive toolbox

Table 1 lists 10 heuristics that are likely to be in the adaptive toolbox of humans. Each of these heuristics appears to violate some standard of rational decision making. Consider strategies that make no “trade-offs” between cues (for this, see Sect. 4), such as take-the-best and elimination-by-aspect (Tversky 1972), and that are called noncompensatory strategies. For instance, a review of 45 process-tracking studies

concluded that when people choose between alternatives, from apartments to contraceptive pills, there is overwhelming evidence for noncompensatory heuristics, whereas evidence for weighting-and-adding strategies is restricted to tasks with small numbers of alternatives and attributes (Ford et al. 1989). Recent studies have provided further evidence for heuristics that make no trade-offs in consumer choice (Kohli and Jedidi 2007; Yee et al. 2007). Studies also showed that when people make inferences, their reaction times and behavior can be predicted by take-the-best and similar heuristics typically better than by competing strategies, in particular in situations when relying on one good reason is adaptive (Bergert and Nosofsky 2007; Bröder and Gaissmaier 2007; Nosofsky and Bergert 2007; Rieskamp and Otto 2006). In one area, however, evidence remains scarce, in part because of the complete dominance of weighting-and-adding theories. In choices between gambles, the evidence for one-good-reason heuristics such as the priority heuristic (Brandstätter et al. 2006) is under debate (e.g., Brandstätter et al. 2008; Johnson et al. 2008). Note that not all heuristics are non-compensatory. Tallying, for instance, makes trade-offs. This is a strategy named after the tally sticks for counting, which can be traced back some 30,000 years in human history. It uses equal weights (as in $1/N$) and can predict almost as accurately as—and often better than—multiple linear regression (see Dawes and Corrigan 1974; Dawes 1979; Einhorn and Hogarth 1975; Czerlinski et al. 1999; Schmidt 1971).

The toolbox of heuristics is described as “adaptive” in order to express the idea that our reasoning strategies have an evolutionary background. Human beings, therefore, share at least some heuristics with other animals. The degree to which species share heuristics will depend on whether they share the evolved capacities on which a heuristic feeds. For instance, humans are genetically disposed to produce language and provided with cultural input for learning a particular language. It is through language that legal and social defaults can be communicated, for instance, that by French law every citizen is a potential organ donor, unless he or she opts out, whereas in the UK no citizen is a donor, unless he or she opts in. Since most citizens in both countries rely on the same default heuristic (“if there is a default, do nothing”, see Table 1), the result are strikingly different rates of organ donors, 99.9% in France compared to only 16% in the UK (Johnson and Goldstein 2003). A heuristic such as the default heuristic requires evolved capacities that are unique to humans, and are thus not shared with other animals.

Whereas language production is not in the adaptive toolbox of animals, or only in very rudimentary forms, we share other core capacities with many species, such as object tracking and recognition memory. A shared capacity between two species makes it likely that these rely on similar heuristics, even if they have to solve different problems, such as intercepting prey as opposed to fly balls. If the problem is the same but evolved capacities do not match, different heuristics are required. Consider estimation of area. Ants, for instance, can produce pheromone trails. To measure the area of a candidate nest cavity, a narrow crack in a rock, an ant has no yardstick but a rule of thumb based on this capacity: “Run around on an irregular path for a fixed period, and lay down a pheromone trail. Then leave. Return to the cavity, move around on a different irregular path, and estimate its size by the frequency of encountering the old trail.” This heuristic is remarkably precise: Nests half the area of others yielded reencounter frequencies 1.96 times greater (Mugford et al. 2001). Many of these evolved rules

Table 1 Ten well-studied heuristics for which there is evidence that they are in humans' adaptive toolbox

Heuristic	Definition ^a	Ecologically rational if:	Surprising findings (examples)
Recognition heuristic (Goldstein and Gigerenzer 2002)	If one of two alternatives is recognized, infer that it has the higher value on the criterion	Recognition validity >0.5	Less-is-more effect if $\alpha > \beta$; systematic forgetting can be beneficial (Schooler and Hertwig 2005). Can beat Wimbledon experts and ATP rankings in predicting outcomes of Wimbledon matches (Serwe and Frings 2006)
Fluency heuristic (Schooler and Hertwig 2005)	If both alternatives are recognized but one is recognized faster, infer that it has the higher value on the criterion	Fluency validity >0.5	Less-is-more effect; systematic forgetting can be beneficial
Take-the-best (Gigerenzer and Goldstein 1996)	To infer which of two alternatives has the higher value: (1) search through cues in order of validity, (2) stop search as soon as a cue discriminates, (3) choose the alternative this cue favors	Cue validities vary moderately to highly, moderate to high redundancy	Often predicts more accurately than multiple regression (Czerlinski et al. 1999), neural networks, exemplar models, and decision tree algorithms (Brighton 2006)
Tallying (unit-weight linear model, Dawes 1979)	To estimate a criterion, do not estimate weights but simply count the number of positive cues	Cue validities vary little, low redundancy (Hogarth and Karelaia 2005, 2006)	Often predicts as accurately as or more accurately than multiple regression (Czerlinski et al. 1999)
Satisficing (Simon 1955; Todd and Miller 1999)	Search through alternatives and choose the first one that exceeds your aspiration level	Number of alternatives decreases rapidly over time, such as in seasonal mating pools (Dudey and Todd 2002)	Aspiration levels can lead to substantial better choice than chance, even if they are arbitrary [e.g., the secretary problem, see Gilbert and Mosteller (1966); the envelope problem, see Bruss (2000)]
1/N; equality heuristic (DeMiguel et al. 2009; Messick 1993)	Allocate resources equally to each of N alternatives	High unpredictability, small learning sample, large N	Can outperform optimal asset allocation portfolios

Table 1 continued

Heuristic	Definition ^a	Ecologically rational if:	Surprising findings (examples)
Default heuristic (Johnson and Goldstein 2003)	If there is a default, do nothing	Values of those who set defaults match those of the decision maker; when the consequences of a choice are hard to foresee	Explains why mass mailing has little effect on organ donor registration; predicts behavior when trait and preference theories fail
Tit-for-tat (Axelrod 1984)	Cooperate first and then imitate your partner's last behavior	The other players also play tit-for-tat; the rules of the game allow only defection or cooperation, but not divorce	Can lead to a higher payoff than optimization (backward induction)
Imitate the majority (Boyd and Richerson 2005)	Consider the majority of people in your peer group and imitate their behavior	Environment is stable or only changes slowly; info search is costly or time-consuming	A driving force in bonding, group identification, and moral behavior
Imitate the successful (Boyd and Richerson 2005)	Consider the most successful person and imitate his or her behavior	Individual learning is slow; information search is costly or time-consuming	A driving force in cultural evolution

^a For formal definitions, see references. Each heuristic used to solve problems in social and nonsocial environments. The conditions concerning the ecological rationality and predictions are explained in the references given

of thumb in animals, some of which can be found in humans as well, are amazingly simple and efficient (see the overview by [Hutchinson and Gigerenzer 2005](#)). Via their building blocks based on evolved abilities, they process the patterns of information available from the environment to produce goal-directed behavior.

2.2 Building blocks of heuristics

Most heuristics consist of several building blocks and can be adapted to new problems by having one or more of their building blocks substituted. The number of building blocks is consequently smaller than the number of heuristics they can create. Building blocks include search rules, stopping rules, and decision rules. For instance, to choose a mate, a peahen does not investigate all peacocks posing and displaying in a lek eager to get her attention, nor does she weight and add all male features to calculate the one with the highest expected utility. Rather, she investigates only three or four and selects the one with the largest number of eyespots ([Petrie and Halliday 1994](#)). This corresponds to a simple search rule (*investigate males in your proximity*), a stopping rule (*stop search after a sample of four*), and a decision rule that relies on one good reason only.

2.3 Evolved capacities

The building blocks of heuristics, as mentioned above, are based on evolved capacities. An evolved capacity is a product of nature and nurture, a capacity that is prepared by a species' genes but needs practice to express itself. For instance, 3-month-old babies spontaneously practice holding their gaze on moving targets, such as a mobile; this capacity is later necessary for using the gaze heuristic to catch balls in baseball or cricket: "Fix your gaze on the ball, start running, and adjust your running speed so that the angle of gaze remains constant" ([Gigerenzer 2007](#)). Evolved capacities are one reason why simple heuristics can perform so well: They enable finding solutions to complex problems that are fundamentally different from the mathematically inspired idea that humans and animals somehow estimate trajectories.

2.4 How are heuristics selected?

In their seminal work on adaptive decision making, [Payne et al. \(1993\)](#) collected evidence that people tend to select heuristics in an adaptive way. This evidence focused on preferential choice. Similar evidence was obtained for the adaptive use of heuristics in inductive inference (e.g., [Bröder 2003](#); [Rieskamp and Hoffrage 2008](#); [Dieckmann and Rieskamp 2007](#)). The fact that people tend to rely on heuristics in situations where these are ecologically rational raises a new question. How does the mind select heuristics from the adaptive toolbox? This process is mostly unconscious and only partly understood, but we know of three selection principles.

2.4.1 Memory constrains selection

First, consider the choice between the first three heuristics in Table 1: the recognition heuristic, the fluency heuristic, and take-the-best. Consider also the 2003 Wimbledon Gentlemen's tennis match between Andy Roddick and Tommy Robredo. What strategy can a spectator use to place a bet on who will win before the start of the game? If one has heard of Roddick but not of Robredo, then this state of memory restricts the choice set to the recognition heuristic (which leads to a correct prediction: Roddick won); if one has heard of both but knows nothing except their names, that restricts the choice to the fluency heuristic; and if one has heard of both and knows some additional facts, then there is a choice between the fluency heuristic and take-the-best. If none of the names are available in memory, then none of these three heuristics applies. That does not mean that one needs to guess; instead it is possible to check only the betting quotient and imitate the majority, that is, bet on the player on whom most others bet.

2.4.2 Learning by feedback

The second selection principle is feedback, which applies to the set of heuristics that remain after memory constraints. Strategy selection theory (Rieskamp and Otto 2006) provides a quantitative model that can be understood as a reinforcement theory where the unit of reinforcement is not a behavior, but a heuristic. This model facilitates predictions about the probability that a person selects one strategy within a defined set of strategies.

2.4.3 Environmental selection

The third selection principle relies on the structure of the environment. For instance, the recognition heuristic is likely to lead to accurate (and fast) judgments if the recognition validity is high, that is, if a strong correlation between recognition and the criterion exists, as is the case for tennis players and the chance that they will win the game. There is experimental evidence that people tend to rely on this heuristic if the recognition validity (called α) is high, but less so if it is low or at chance level ($\alpha = 0.5$). For instance, name recognition of Swiss cities is a valid predictor for their population ($\alpha = 0.86$), but not for their distance from the center of Switzerland, the city of Interlaken ($\alpha = 0.51$). It was reported that 89% of participants relied on the recognition heuristic in judgments of population, but only 54% in judgments of distance to Interlaken (Pohl 2006). Thus, the same participants changed their reliance on the recognition heuristic when judging the same cities, depending on the correlation between recognition and the criterion.

2.5 What is not a heuristic?

Strategies such as the mean-variance portfolio are not of heuristic nature. They attempt to weight and add all available information and make use of heavy computation. The origins of these optimization theories can be traced back to the theory of rationality

that emerged during the Enlightenment (Daston 1988; Gigerenzer et al. 1989). Benjamin Franklin described the ideal of weighting and adding all reasons in a letter to his nephew:

April 8, 1779

If you doubt, set down all the Reasons, pro and con, in opposite Columns on a Sheet of Paper, and when you have considered them two or three Days, perform an Operation similar to that in some questions of Algebra; observe what Reasons or Motives in each Column are equal in weight, one to one, one to two, two to three, or the like, and when you have struck out from both Sides all the Equalities, you will see in which column remains the Balance. [...] This kind of *Moral Algebra* I have often practiced in important and dubious Concerns, and tho' it cannot be mathematically exact, I have found it extremely useful. By the way, if you do not learn it, I apprehend you will never be married.

I am ever your affectionate Uncle,
B. Franklin

This vision has inspired the view that complex problems need to be solved by complex algorithms, and that more information is always better. Carnap (1947) proposed the “principle of total evidence,” which is the recommendation to use all the available evidence when estimating a probability. Similarly, the statistician Good (1967) argued that it is irrational to leave observations in the record but not use them. Modern versions of Franklin’s moral algebra include expected utility maximization in economics, Bayesian inference theories in the cognitive sciences, Markowitz’s mean-variance optimization model in finance, various bookkeeping principles taught in MBA courses and recommended by consulting firms, and much of the theories of rationality taught in philosophy curricula.

Note that Franklin ends with the advice that learning his moral algebra is vital for marriage. To check, we asked a sample of economists who teach the modern versions of this theory whether they had chosen their partner using their favorite method. Only one said that he had. He explained that he listed all the options available and all the important consequences that he could think of, such as whether she would remain interesting to talk to after the honeymoon excitement was over, take care of the children, and let him work in peace. He took several days time to estimate the utilities of each of these consequences, as well as the probabilities that they might actually occur for each woman. After calculating the expected utility for each candidate, he proposed to the woman with the highest value—without telling her how he had made his choice—and she accepted. They are now divorced.

The point of this story is emphatically not that Franklin’s bookkeeping method produces more divorces than a simple imitate-the-majority heuristic (Table 1): *Try*

to get the mate that your peers desire. Rather, our point is that there is a discrepancy between normative theory and actual practice: Despite its being advertised as *the* rational method, even devoted proponents often instead rely on heuristics in important decisions. Markowitz is no exception. Reinhard Selten, another Nobel laureate in economics, who made the “chain store problem” famous by proving that an aggressive policy against competitors is futile, found his logically correct proof intuitively unconvincing and said that in the real world, he would rather follow his gut feeling to be aggressive (Gigerenzer 2007). Many philosophers claim that we should view norms of rationality as being justified by what Goodman (1965) and Rawls (1971) have called the “reflective equilibrium” between principles and intuitions: Whereas norms can be derived neither from mere descriptions of human behavior (or particular intuitions about what to do) nor from some metaphysical cloud-cuckoo land, normative theories can be constructed by striking a balance between a comprehensive amount of intuitions and principles. To strengthen the equilibrium even more, the experts’ best theories should be included in it (Daniels 1979). But as the case of the two Noble laureates shows, experts often do not follow their own theories, and for a good reason: Any theory of rationality is ecologically rational in specific environments, not in all possible worlds.

Health is another example in point. In a study, more than 100 economists were asked how they decided whether or not to take a prostate-specific antigen (PSA) test (Berg et al. 2011). Virtually all medical societies recommend that patients carefully weight pros and cons of PSA testing, since its benefit remains controversial (it is not proven that it saves lives) but its harms are documented (such as incontinence and impotence from operations following positive tests). Yet two thirds of the economists said that they had not weighted any pros and cons but did whatever their doctors or their wives said they should do. Here, the social heuristic *trust your doctor* guides important decisions. Note that our point here is descriptive, not evaluative: We question whether the bookkeeping method, understood as a norm, is justified and contend that it is neither good nor bad per se. In terms of the normative question, we propose to ask in what environment a given strategy (such as moral algebra) is better. For instance, in a world where doctors practice defensive decision making for fear of malpractice lawsuits and where most doctors do not have the time and education to read the relevant medical studies but are “informed” by representatives from the pharmaceutical industry, it pays to carefully weigh benefits and risks rather than rely on the trust-your-doctor heuristic (Gigerenzer et al. 2007).

3 Heuristics and the environment: an ecological conception of rationality

We now take a step closer towards the normative issues. To do so, we use the concept of “ecological rationality” in order to indicate that heuristics (like other methods of inference, including optimization) are to be evaluated in terms of their relation to given environments in which humans reason and decide. Following some preliminary remarks about the concept of rationality proposed here, we then explain the notion of environment. Afterwards, we make clear why looking at the relation of heuristics to environments is often normatively more useful than evaluating reasoning and decision

making according to standard norms of probability or decision theory that aim at an optimum of information and computation.

3.1 Ecological rationality

The concept of ecological rationality is inextricably linked to that of the adaptive toolbox. Cognition in an uncertain world would be inferior, inflexible, and inefficient with a general-purpose calculator (Gigerenzer and Brighton 2009). For instance, nocturnal moths migrate using the moon as primary reference point, but on moonless nights they navigate solely with a geomagnetic compass. Humans can solve the same problem by drawing on an adaptive toolbox that includes many social heuristics, such as asking others for directions or just following the crowd. A handyman knows that there is no single best tool for every task, a universal hammer, but a box of tools that need to be used in an adaptive way.

We use the term *ecological rationality* for both a specific normative conception of rationality and a specific research program. As a normative conception, it provides an alternative to views of rationality that focus too exclusively on internal consistency, coherence, or logic as opposed to the relationship between mind and environment. Ecological rationality is about the success of cognitive strategies in the world, as measured by currencies such as the accuracy, frugality, or speed of decisions. This term was introduced to flesh out Herbert Simon's adaptive view of rational behavior (Gigerenzer et al. 1999). As Simon put it, "Human rational behavior ... is shaped by a scissors whose two blades are the structure of the task environments and the computational capabilities of the actor" (Simon 1990, p. 7). Shortly before his death, Simon responded that the ecological approach brings "sanity into the study of rationality" (see Gigerenzer 2004), and Vernon Smith used the same concept in the title of his Noble Laureate lecture (Smith 2003). Yet to date, this conception remains a small island in the Holy Roman Empire of normative theories of rationality.

The study of ecological rationality, understood normatively, poses two questions:

Given a strategy, in what environments does it succeed?

Given an environment, what strategies succeed in it?

The investment example provided answers to both questions, which are intimately related. For instance, given $1/N$, investment environments with large N and a relatively small sample size are the right match. Or, given an environment with $N = 50$ and 10 years of stock data, $1/N$ is likely to perform better than the mean-variance portfolio. The study of ecological rationality uses proofs or computer simulation to find out when ignoring what information is rational. Table 1 provides further results.

Note that while ecological rationality is broadly defined in terms of success, and thus involves looking for means suited to certain goals, we do not maintain that reasoning is only about satisfying desires, without caring what is actually true or correct. For instance, the take-the-best heuristic has, on average, made correct predictions more often than optimization strategies (Czerlinski et al. 1999; Gigerenzer and Brighton 2009). We thus measure the success of this heuristic against standard epistemic goals.

3.2 The structure of environments

An *environment* is what an agent acts in and upon. The environment also influences the agent's actions in multiple ways, by determining the goals that the agent aims to fulfill, shaping the tools that the agent has for reaching those goals, and providing the input processed by the agent. We do not yet know whether it is possible to completely classify environmental structures, but several important structures have been identified. The analysis of the investment problem revealed three environmental structures: the degree of uncertainty, the number of alternatives, and the size of the learning sample. These are of relevance for a number of tasks.

3.2.1 *Uncertainty*

The degree of uncertainty refers to the extent that available cues can predict a criterion. Predictability is sometimes measured by R^2 , but other measures exist. Uncertainty varies by the kind of criterion and prediction. Tomorrow's performance of stocks and funds is highly unpredictable, heart attacks are slightly more predictable, and tomorrow's weather is the most predictable of these three criteria. Furthermore, uncertainty is higher when one has to make out-of-population rather than out-of-sample predictions (Gigerenzer and Brighton 2009). The investment example illustrates that the higher the uncertainty, the greater the advantage of simple heuristics compared to optimization methods. There is an easy intuitive way to understand this result. In a world without uncertainty—the world Laplace's superintelligence inhabits—all relevant past information needs to be considered. In a fully unpredictable world, such as a perfect roulette wheel, one can ignore all information about past performance. But most of the time, humble humans live in the twilight world of in-between certainty, where the general lesson is that the lower the predictability of the criterion, the more information they should ignore, relying instead on a good heuristic principle. $1/N$, for instance, ignores all information about the previous performance of funds and relies on the simplest of all diversification principles.

3.2.2 *Number of alternatives*

In general, problems with large number of alternatives pose difficulties to optimization methods. *Alternatives* can refer to individual objects (such as funds) or sequences (such as moves or pathways). Problems with a large number of possible moves, such as chess, do not allow for optimization. Although an optimal (best) sequence of moves does exist in chess, no computer or mind can determine it, since the problem is *computationally intractable*. Deep Blue and human chess masters alike have to rely on nooptimization techniques, including heuristics. Similarly, the best route in the traveling salesman problem and the best strategies in the computer games Tetris and Mine-sweeper are computationally intractable (Michalewicz and Fogel 2000). In general, an optimization model defines a problem by making a number of assumptions that facilitate calculations and then proving that a strategy exists for optimizing a criterion. There exist well-defined problems, such as tic-tac-toe, where one can prove the optimal strategy. But once this strategy is known, the game becomes boring.

The key disadvantage to the routine use of optimization techniques is that the problem in question often needs to be replaced by a mathematically convenient small-world problem (Savage 1972) that can be optimized. For instance, if building an optimization model for chess, one would have to make it tractable in the first place, for instance, by shrinking the 8×8 board into a 4×4 board, with a correspondingly reduced set of figures. Whether the optimal strategy found in the miniature game is also the optimal one in the real problem, however, remains open. As the case of the mean-variance portfolio illustrates, an optimization model does not imply optimal behavior in the real world.

3.2.3 Sample size

In general, the smaller the sample size, the greater the advantage for simple heuristics. One of the reasons is that complex statistical models have to estimate their parameter values from past data, and if the sample size is small, then the resulting error due to “variance” may exceed the error due to a heuristic’s “bias” (Gigerenzer and Brighton 2009). What a small sample size is depends on the degree of uncertainty, as can be seen in the investment problem where uncertainty is high and a sample size of hundreds of years of stock data is needed to match and surpass the accuracy of $1/N$.

There are also other relevant structures of environments, for instance, redundancy and variability:

Redundancy Redundancy of cues is often measured by the average correlation between the cues. Redundant cues can be exploited by sequential heuristics such as take-the-best that rely on the first good reason that allows a decision to be made (Dieckmann and Rieskamp 2007).

Variability The variability of cue weights can be exploited by heuristics that stop search quickly. For instance, when variability is high, heuristics that rely on the best cue alone perform better than when variability is low (Martignon and Hoffrage 2002; Hogarth and Karelaia 2005, 2006).

Note that our use of the term *environment* is not identical with the physical or “real” environment (Todd 2001). For instance, the degree of uncertainty reflects the environment (*ontic uncertainty*) as well as the mind’s limited understanding (*epistemic uncertainty*); hence, the degree of uncertainty is located in the system mind–environment. Similarly, the number of alternatives and the sample size depend on both what is available in an environment and what an actor or investor actually included in the consideration set (such as the number N of funds in the analysis). Finally, redundancy and variability of cues depend both on what is in the physical environment and on humans who select certain cues and not others, which can result in more or less redundant cues. Physicians, for instance, tend to look at redundant cues, but they could choose to analyze independent cues and thus in part create their environment.

As important as physical or biological environments are social environments. For instance, although an investment can be decided upon individually and made on the Internet without interacting with other people, the stock market itself is driven by both nature (e.g., a disastrous hurricane) and people (e.g., the advent of securitization, that is, the bundling and selling of mortgages). Each of the heuristics in Table 1 can be used

for social objects (e.g., whom to hire, to trust, to marry) as well as for physical objects (e.g., what goods to buy). For example, the recognition heuristic exploits environmental structures in which lack of recognition is valuable information and aids inferences about whom to talk to and whom to trust (*don't talk to strangers*), but also about in which stocks to invest. Similarly, satisficing can assist in choosing a mate (Todd and Miller 1999), but also in buying a pair of trousers, and $1/N$ can help parents to allocate their love and time equally to their children, creating a sense of fairness, but also help investors to diversify.

4 The normative issue: why should people rely on heuristics?

Finally, what are the reasons in favor of using heuristics over optimization strategies from a normative point of view? To answer this question, we provide critical considerations of why optimization strategies are often inadequate guides for problem solving and give two reasons in favor of heuristics. This will also reveal that there are (at least) two different ways of naturalizing the kind of rationality under discussion here. We begin with experts.

4.1 What strategies do experts rely on?

A first argument against the rationality of simple heuristics is based on widespread assumptions about the behavior of experts. They are believed to always use optimizing strategies; how could they otherwise be counted as experts? As in the case of Markowitz and the other examples given in Sect. 2, empirical studies have shown that experts typically know to distinguish between a heavenly optimization model and what works in the real world. In general, experts use less information than novices; they know where to look and tend to rely on limited search and simple heuristics (Ettenson et al. 1987; Dhimi 2003; Dhimi and Aytton 2001; Shanteau 1992). This is illustrated by a study on burglary in which graduate students were given pairs of residential properties described by eight binary cues, such as flat/house, letter box empty/stuffed with post, and burglar alarm system yes/no (Garcia-Retamero and Dhimi 2009). The students were asked which property was more likely to be burgled. Two models of cognitive processes were tested, weighting-and-adding of all information and take-the-best, a simple lexicographic heuristic that bases its decision on the most important cue only (Table 1). The result was that 78% of the students were classified as relying on weighting-and-adding and 22% on take-the-best. Normally, psychological experiments stop here. Yet the authors then studied experts, namely convicted burglars from a male English prison, who reported having successfully committed burglary on an average of 57 occasions. Burglars' decisions were strikingly different: 72% could be predicted by take-the-best and 28% by weighting-and-adding. A second expert group, police officers who had investigated residential burglaries, showed similar reliance on take-the-best. These findings are consistent with other studies concluding that experienced experts such as managers and physicians tend to rely on simple heuristics (Gigerenzer 2007; Wübben and Wangenheim 2008).

4.2 But why not optimization?

Although optimization is considered the sine qua non in many theories of rational choice, its relevance for decisions in the real world is limited. Here are the major reasons why optimization models are rarely of use to experts and laypeople alike.

4.2.1 Tractability

Most interesting real-world problems are computationally intractable, that is, they lead to combinatorial explosion (Reddy 1988). For instance, probabilistic inferences using Bayesian belief networks (Cooper 1990), including approximations (Dagum and Luby 1993), are NP-hard. Once a problem is deemed intractable, it is appropriate to direct efforts to finding out how people nevertheless solve it, rather than to reducing the problem to a tractable size in order to allow for optimization.

4.2.2 Robustness

Consider now the subset of problems that are tractable, such as computing a mean-variance portfolio, where the optimization calculations can be performed. Yet optimization, as we have seen, was inferior to $1/N$ unless one had hundreds of years of stock data. The main reason is that the estimates of the parameters were not robust, that is, they suffered from “variance” (Gigerenzer and Brighton 2009), due to the high uncertainty and relatively small sample sizes in the environment. In general, optimization can only lead to optimal outcomes if its parameters are estimated (almost) without error, for which one requires environments with low uncertainty and large sample size. Robustness is a main reason why simple heuristics can lead to more accurate inferences than sophisticated statistical software programs can.

4.2.3 Ill-defined problems

Unlike chess and well-defined experimental games, real-world environments are often ill-defined. The necessary ingredients for optimization—such as the complete set of alternatives, consequences, utilities, and probability distributions—are only partially known. In the investment problem, all parameters are estimated, but in other situations they are not even known. There is a further distinction between naturally and intentionally ill-defined situations. How to find the ideal husband, the best job, or the meaning of life are naturally ill-defined problems. We do not know the set of alternatives, probabilities, and so on, and if we did, these problems would cease to be the big questions in life. Some legal contracts are examples of intentionally ill-defined situations. For instance, marriage contracts often intentionally do not specify every consequence for every undesirable action, in order to retain some degree of decency and trust. Although one could write a better-specified contract, neither party is likely to want this. In such situations, the calculating spirit of the optimization method, as illustrated in Franklin’s moral algebra, is considered inappropriate or even morally reprehensible. If a person chooses a spouse by calculation, and he or she finds out, the spouse may well react with revulsion. In contrast, falling in love, even if one cannot

say why or simply desires the person that everyone else admired, is the expected basis of choice for young Western men and women.

4.2.4 Novelty and surprise

An optimization method typically optimizes on past events, and if the future is not like the past, the method's usefulness is limited in yet another way. We distinguish three types of uncertainty in prediction. The first is out-of-sample prediction, where one knows a sample of events in a population, and one has to make predictions about another sample. This corresponds to the investment problem, where the performance of funds up to a particular time is known, and one has to predict their performance in the next month, assuming the market is stable. With this first kind of uncertainty, optimization methods can already fail to produce better results than a simple heuristic, depending on the conditions. The second type of uncertainty incurs in out-of-population prediction, where one has information about one population and needs to draw conclusions about another population that differs in unknown ways. For instance, when a diagnostic system for heart attacks is validated on a sample of patients in Boston and applied to patients in Michigan, we face out-of-population prediction. Here, robustness becomes even more important than in out-of-sample prediction and can be reached by radically simplifying a diagnostic system, such as replacing a logistic regression with fast-and-frugal trees (Green and Mehr 1997; Martignon et al. 2003). The third type of uncertainty is what we call novelty or surprise, where unlike in the other two cases, all alternatives and consequences are not known. For instance, due to global warming, more and more animal species living in the northern hemisphere are surprised by new predators moving north. That is, unforeseen players emerge. To be prepared for novelty, coarse behavior that appears to be rigid and inflexible has been shown to be superior to behavior that is fine-tuned and optimized to a past environment that is assumed to be eternally stable (Bookstaber and Langsam 1985). In other words, to be prepared for novelty, humans and other animals need to deviate from the optimality conditions and show behavior that has a chance of surviving when environments change.

To summarize, contrary to the widespread use of optimization in theory (as opposed to actual expert practice), there are several good reasons not to rely routinely on this technique for understanding the “oughts” and “is” of human behavior (Gigerenzer 2004; Selten 2001).

4.3 Why should people rely on heuristics? Two kinds of arguments (and heuristics)

The classical justification for heuristics is that people save effort, albeit at the cost of accuracy. There is excellent work showing that, depending on the structure of the environment, the loss of accuracy can be small (Payne et al. 1993). This interpretation of the function of heuristics is known as:

Accuracy–effort trade-off Humans and other animals rely on heuristics because information search and computation cost time and effort; thus, they trade off some loss in accuracy for faster and more frugal cognition.

This view starts from the dictum that more is always better, but since information is not free in the real world, so the argument goes, there is a point at which the costs of further search exceed the benefits. It underlies optimization-under-constraints theories in which search in the world (e.g., [Stigler 1961](#)) or in memory (e.g., [Anderson 1990](#)) is terminated when the expected costs exceed its benefits. Similarly, the analysis of the adaptive decision maker ([Payne et al. 1993](#)) is built around the assumption that heuristics achieve a trade-off between accuracy and effort, where effort is a function of the amount of information and computation consumed. And indeed it is correct that heuristics often save effort. But with this interpretation, we do not need to revise the normative assumption that more information and computation are better than less.

The major discovery of recent research on heuristics, however, is that saving effort does not always lead to a loss in accuracy. Heuristics can be faster and more accurate than strategies that use more information and more computation. Our analysis of the function of heuristics goes beyond accuracy–effort trade-offs:

Less can be more Humans and other animals rely on heuristics in situations where these are ecologically rational, including situations where less information and computation lead to more accurate judgments.

Here is the essential difference. The interpretation that relying on a heuristic is functional because of an accuracy–effort trade-off is based on gold standards of rationality, including weighting-and-adding models in the spirit of Franklin’s moral algebra and Bayes’ rule. By its very definition, no heuristic can be more accurate. In contrast, the study of ecological rationality treats the normative question as an empirical one and allows and investigates *less-is-more effects*, as in the investment example.

Less-is-more effects are surprising indeed. One explanation for them is that a strategy can make two kinds of errors: *bias* and *variance*. Specialized tools such as heuristics tend to have larger bias than more general ones, and general-purpose tools with many free parameters tend to generate more variance. A good cognitive system needs strategies that strike a balance between being too specialized and being too general. From this it follows that a mind with a general-purpose algorithm would not only be slow, but also perform inferior to a more specialized system if the error due to variance was larger than the error due to bias (for details, see [Brighton and Gigerenzer 2008](#); [Gigerenzer and Brighton 2009](#)).

Less-is-more effects have been studied and demonstrated for some time, such as in the seminal work by Dawes and colleagues on tallying ([Dawes and Corrigan 1974](#); [Dawes 1979](#)) and on take-the-best ([Gigerenzer and Brighton 2009](#); [Gigerenzer and Goldstein 1996](#)). Yet relevant studies have been ignored for decades ([Makridakis and Hibon 2000](#)). When issues of rationality are at stake, emotions can arise that are resistant to evidence. If you open a book on cognition or decision making, the following three misconceptions are still likely to be asserted:

Misapprehension 1 Heuristics are second-best; optimization is always better.

Misapprehension 2 We use heuristics only because of our cognitive limitations.

Misapprehension 3 More information, more computation, and more time are always better.

Each statement stems from the assumed generality of the accuracy–effort trade-off. Yet each is incorrect as a general belief, and more and more studies document less-is-more effects. For instance, in an age in which companies maintain databases of their customers, complete with historical purchase data, a key question becomes predicting which customers are likely to purchase again in a given time frame, and which are inactive. Wübben and Wangenheim (2008) found that managers in airline and apparel industries rely on a simple hiatus heuristic: “If a customer has not purchased within 9 months in the past (the ‘hiatus’), the customer is classified as inactive, and otherwise active”. They compared the hiatus heuristic with the Pareto/NBD model, which assumes that purchases follow a Poisson process with a purchase rate parameter λ , that customer lifetimes follow an exponential distribution with a dropout rate parameter μ , and that, across customers, purchase and dropout rates are distributed according to a gamma distribution. For both industries, the simple heuristic correctly classified more customers than the Pareto/NBD model. Similarly, in library search, a one-reason decision heuristic produced better order of titles than a Bayesian model and PsychInfo (Lee et al. 2002). For many inference problems in the real world, there is an inversely U-shaped relation between amount of information, computation, and time on the one hand and accuracy on the other.

There are thus two quite different reasons for why people do—and ought to—rely on heuristics: For one, there are many tasks where a trade-off is made between accuracy and effort. At the same time, there are also many tasks where a less-is-more effect is visible. In the first of these cases, heuristics are to be recommended in order to save time or energy, thereby leaving open whether there are other norms that, given only more time or energy, would lead to more accurate or better results. In the second case, however, it would be irrational to apply norms that require more information or computation. Psychology can thus be used to derive norms in two different ways: first, in order to save time or other resources, and second, in order to refute the widespread universal claim that more information and computation always lead to better results. Experiments show that they do not.

5 An objection? On the distinction between the justification and application of norms

According to our normative naturalism, people should often rely on heuristics. There is an objection to this, however, that starts from a plausible distinction: It is one thing to justify a particular norm or rule and another thing to apply it correctly, that is, to recognize the situations in which it should be applied (see, e.g., Alexy 1993; Günther 1993; Habermas 1997). If these are not properly recognized, mistakes or dogmatism about norms can ensue. A tax law may be perfectly justified within a certain system of laws, yet it may not apply to certain persons—for instance, foreigners who pay their taxes elsewhere. Accordingly, we blame or change the judge, not the law, if he or she incorrectly subsumes a certain case under it. Although a logical rule may be perfectly justified within a particular calculus—and even derivable within several logical systems—it may be inappropriate when applied to certain cases of ordinary reasoning. For instance, the conjunction rule, which states that the probability of a conjunction of

events cannot exceed the probability of each conjunct, is a justified truth of the probability calculus. Yet, it may be inappropriate to apply it as a norm in certain reasoning tasks (as discussions of the notorious “Linda problem” have clarified; see [Hertwig and Gigerenzer 1999](#), with reference to Grice et al.’s [1975] conversational “maxim of relevance”). More generally, formal rules cannot simply be equated with norms of rationality ([Goldman 1986, 2008](#)).

Now, the objection continues, if a particular law or rule should not be applied in certain situations, this hardly implies that the law or rule in question is not perfectly justified or that ideal reasoners should not use it. A variety of such thinking occurs in ethical discussion about utilitarianism: It may be difficult to identify the rule or action that is optimal in the sense that it maximizes happiness or expected utility, yet that does not undermine the highest principle of utilitarianism. We can distinguish between “practical” and “ideal” norms ([Birnbacher 1988](#)). It is obvious how such an argument might be transferred to the debate about the normative evaluation of heuristics, at least those that are based on accuracy–effort trade-offs: They fall short of some optimization strategy, which itself may be justified by a comprehensive normative theory of rationality. There are various ways in which the objection can be formulated, related to different meanings and uses of the term ‘rational’, some of which we hinted at in the introduction. For instance, one might say that such heuristics are normatively required only in practice, not in principle. Alternatively, one might say that we need to distinguish between the *rationality of rules* and the *rationality of agents* or their inferences and decisions. According to the objection starting from the distinction between justification and application, we are confusing the question of whether a norm is justified with the question of the case-sensitive application of the norm. Alternatively, our position trades on ambiguities in the concept of rationality. Does that not undermine our naturalization of rationality?

We see two main problems with this objection. First, the argument delivers merely a promise. It has to be shown in each and every case where a trade-off between accuracy and effort is made that there is indeed an optimization strategy that is theoretically better justified. And even if such proof were given, there would often be good, namely practical reasons for not using the optimization strategy in the real world. Second, and more importantly, the argument does not work for less-is-more effects. Here, a simple heuristic can be equally or more accurate than a strategy that relies on more information and computation, including optimization methods. Hence, no formal system, from versions of logic to those of Bayesianism, can be justified as the rational solution for all problems. [Savage \(1972\)](#), who is credited as the father of modern Bayesian decision theory, emphasized that this theory can only be justified as a theory of rationality in a *small world* where all the relevant pieces of information are known and where no surprises occur. Yet contemporary Bayesianism often appears to assume that Bayesian decision theory is a general-purpose tool for rationality in both small and large worlds (see [Binmore 2009](#)). More generally, the application of every strategy, optimization or heuristic, can be justified only relative to the structure of the environment. In our view, optimization and heuristic strategies should therefore be treated equally. Since less-is-more effects exist, the usual distinction between justification and application, which may have been legitimate elsewhere, breaks down.

There is a historical irony here. The distinction between justification and application stands in a venerable tradition, it being closely related to Immanuel Kant's distinction between the faculties of reason and judgment. Kant has often been viewed as a naïve defender of abstract reason. He was, however, fully aware that the proper application of norms is an important task that cannot be achieved by a priori reasoning. To know when and where and how to correctly apply a certain rule is, as he said, a matter of a well-trained and experienced judgment (*Urteilkraft*). Rather than acquiring it through a priori reasoning, one needs to learn to describe a particular situation correctly and to find the one norm among many to be applied to the situation. Kant even forcefully argued that there cannot be an a priori, strictly general rule guiding the application of rules to problem situations: Any such rule would in turn require another rule guiding its application, which leads to an infinite regress (Kant 1781/1787, pp. A132/B171–A134/B174). This essay is silent about Kant's views on the existence and function of a priori principles either in his theoretical or practical philosophy, because the ones he identified and defended have nothing to do with the norms of optimizing conceptions of rationality, especially in the decision sciences. Nonetheless, Kant was profoundly right in claiming that factual knowledge about the (social and the natural) world must be coupled to norms in order to produce good reasoning. This pertains even to optimization strategies: Their applicability also depends on whether they fit with certain environments. Moreover, when a less-is-more heuristic can help to solve problems for which these optimization strategies were originally designed, the very point of their justification within a certain system becomes unclear.

To sum up: We doubt that the justification-application distinction could be used to undermine our normative naturalism. Like Kant, we emphasize that there is no single general-purpose rule of good reasoning and decision making but that there is a choice between several that requires thinking about the structure of the world.

6 Conclusion: heuristics as descriptive and normative

The classical view of heuristics as lazy shortcuts, falling short of some idealized model of general rationality, grants the study of heuristics only a descriptive role. It draws a strict line between how behavior is and how it should be, with psychology answering the first question but silenced on the second, which is said to be in the territory of normative epistemology or decision theory conceived of as nonempirical enquiries. The empirically grounded study of ecological rationality, in contrast, provides a descriptive and prescriptive role for heuristics. In the right environment, a heuristic *can* be better than an optimization model or other complex strategies. This is not to say that optimization strategies have no role at all, but their role is restricted to well-defined problems with (almost) error-free parameter estimates. In this regard, our naturalism and its concomitant notion of ecological rationality extend themselves even to the correct application of optimization strategies. Mind and environment hang together, and a theory of rationality should respect and explicitly deal with both blades of Simon's scissors.

Acknowledgements The authors are grateful to two anonymous referees for various comments and suggestions. Thomas Sturm’s work on this article was supported by the Spanish Ministry for Science and Innovation, Reference number FFI 2008-01559/FISO.”

References

- Alexy, R. (1993). Justification and application of norms. *Ratio Juris*, 6, 157–170.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Berg, N., Biele, G., & Gigerenzer, G. (2011). Logical consistency and accuracy of beliefs: Survey evidence on health decision-making among economists. Unpublished manuscript.
- Bergert, F. B., & Nosofsky, R. M. (2007). A response–time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 107–129.
- Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.
- Birnbacher, D. (1988). *Verantwortung für zukünftige Generationen*. Stuttgart: Reclam.
- Bishop, M. (2008). Reflections on a normative psychology. In A. Beckermann & S. Walter (Eds.), *Philosophy: Foundations and applications* (pp. 25–262). Paderborn: Mentis.
- Bishop, M., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment*. New York: Oxford University Press.
- Bookstaber, R., & Langsam, J. (1985). On the optimality of coarse behavior rules. *Journal of Theoretical Biology*, 116, 161–193.
- Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. New York: Oxford University Press.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: A process model of risky choice. *Psychological Review*, 113, 409–432.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky choice with heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, and Willemsen (2008) and Rieger and Wang (2008). *Psychological Review*, 115, 281–290.
- Brighton, H. (2006). Robust inference with simple cognitive models. In C. Lebiere & R. Wray (Eds.), *Between a rock and a hard place: Cognitive science principles meet AI-hard problems. Papers from the AAAI Spring symposium* (AAAI Tech. Rep. No. SS-06-03, pp. 17–22). Menlo Park, CA: AAAI Press.
- Brighton, H., & Gigerenzer, G. (2008). Bayesian brains and cognitive mechanisms: Harmony or dissonance? In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 189–208). New York: Oxford University Press.
- Bröder, A. (2003). Decision making with the “adaptive toolbox”: Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology*, 29, 611–625.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin and Review*, 14, 895–900.
- Bruss, F. T. (2000). Der Ungewissheit ein Schnippchen schlagen. *Spektrum der Wissenschaft*, 6, 106–107.
- Carnap, R. (1947). On the application of inductive logic. *Philosophy and Phenomenological Research*, 8, 133–148.
- Chiappe, D. L., & Vervaeke, J. (1997). Fodor, Cherniak, and the naturalization of rationality. *Theory & Psychology*, 7, 799–821.
- Cooper, G. F. (1990). Computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, 42, 393–405.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 97–118). New York: Oxford University Press.
- Dagum, P., & Luby, M. (1993). Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artificial Intelligence*, 60, 141–153.
- Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy*, 76, 256–282.
- Daston, L. J. (1988). *Classical probability in the enlightenment*. Princeton, NJ: Princeton University Press.

- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953.
- Dhami, M. K. (2003). Psychological models of professional decision-making. *Psychological Science*, *14*, 175–180.
- Dhami, M. K., & Ayton, P. (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, *14*, 141–168.
- Dieckmann, A., & Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Memory & Cognition*, *35*, 1801–1813.
- Dudey, T., & Todd, P. M. (2002). Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of Bioeconomics*, *3*, 195–215.
- Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, *13*, 171–192.
- Ettenson, R., Shanteau, J., & Krogstad, J. (1987). Expert judgment: Is more information better? *Psychological Report*, *60*, 227–238.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hulst, B. H., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, *43*, 75–117.
- Garcia-Retamero, R., & Dhami, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin & Review*, *16*, 163–169.
- Gigerenzer, G. (2004). Striking a blow for sanity in theories of rationality. In M. Augier & J. G. March (Eds.), *Models of a man: Essays in memory of Herbert A. Simon* (pp. 389–409). Cambridge, MA: MIT Press.
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Press.
- Gigerenzer, G. (2008). *Rationality for mortals*. New York: Oxford University Press.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*, 1–37.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. W. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53–96.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.
- Gigerenzer, G. & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance. How probability changed science and everyday life*. Cambridge: Cambridge University Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gilbert, J. P., & Mosteller, F. (1966). Recognizing the maximum of a sequence. *American Statistical Association Journal*, *61*, 35–73.
- Goldman, A. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.
- Goldman, A. (2008). Human rationality: Epistemological and psychological perspectives. In A. Beckermann & S. Walter (Eds.), *Philosophy: Foundations and applications* (pp. 230–247). Paderborn: Mentis.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.
- Good, I. J. (1967). On the principle of total evidence. *The British Journal for the Philosophy of Science*, *17*, 319–321.
- Goodman, N. (1965). *Fact, fiction, and forecast*. Indianapolis, IN: Bobbs-Merrill.
- Green, L. A., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *The Journal of Family Practice*, *45*, 219–226.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics*, vol. 3: *Speech acts* (pp. 41–58). New York: Academic Press.

- Günther, K. (1993). *The sense of appropriateness*. Albany, NY: SUNY Press.
- Habermas, J. (1997). *Between facts and norms*. Cambridge: Polity Press.
- Hanna, R. (2006). Rationality and the ethics of logic. *The Journal of Philosophy*, 103, 67–100.
- Hertwig, R., & Gigerenzer, G. (1999). The “conjunction fallacy” revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12, 275–305.
- Hogarth, R. M., & Karelaia, N. (2005). Ignoring information in binary choice with continuous variables: When is less “more”? *Journal of Mathematical Psychology*, 49, 115–124.
- Hogarth, R. M., & Karelaia, N. (2006). Take-the-best and other simple strategies: Why and when they work ‘well’ in binary choice. *Theory and Decision*, 61, 205–249.
- Huberman, G., & Jiang, W. (2006). Offering vs. choice in 401(k) plans: Equity exposure and number of funds. *Journal of Finance*, 61, 763–801.
- Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69, 97–124.
- Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302, 1338–1339.
- Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: A comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, 115, 263–273.
- Kant, I. (1998). In J. Timmermann (Ed.), *Kritik der reinen Vernunft*. Hamburg: Meiner. (Original work published 1781/1787.)
- Kim, J. (1988). What is “naturalized epistemology”? *Philosophical Perspectives*, 2, 381–405.
- Kohli, R., & Jedidi, K. (2007). Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26, 380–399.
- Lee, M. D., Loughlin, N., & Lundberg, I. B. (2002). Applying one reason decision making: The prioritization of literature searches. *Australian Journal of Psychology*, 54, 137–143.
- Makridakis, S., & Hibon M. (2000). The M3-competition: Results, conclusions and implications. *International Journal of Forecasting*, 16, 451–476.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory & Decision*, 52, 29–71.
- Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 189–211). Chichester: Wiley.
- Mele, A. & Rawlings, P. (Eds.). (2004). *The Oxford handbook of rationality*. Oxford: Oxford University Press.
- Messick, D. M. (1993). Equality as a decision heuristic. In B. A. Mellers & J. Baron (Eds.), *Psychological perspectives on justice* (pp. 11–31). New York: Cambridge University Press.
- Michalewicz, Z., & Fogel, D. (2000). *How to solve it: Modern heuristics*. New York: Springer.
- Mugford, S. T., Mallon, E. B., & Franks, N. R. (2001). The accuracy of Buffon’s needle: A rule of thumb used by ants to estimate area. *Behavioral Ecology*, 12, 655–658.
- Nosofsky, R. M., & Bergert, F. B. (2007). Limitations of exemplar models of multi-attribute probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 999–1019.
- Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Petrie, M., & Halliday, T. (1994). Experimental and natural changes in the peacock’s (*Pavo cristatus*) train can affect mating success. *Behavioral and Ecological Sociobiology*, 35, 213–217.
- Pohl, R. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, 19, 251–271.
- Putnam, H. (1982). Why reason can’t be naturalized. *Synthese*, 52, 3–24.
- Quine, W. V. O. (1969). Epistemology naturalized. In W. V. O. Quine (Ed.), *Ontological relativity and other essays* (pp. 69–90). New York: Columbia Press.
- Quine, W. V. O. (1990). *Pursuit of truth*. Cambridge, MA: Harvard University Press (2nd ed. 1992).
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Reddy, R. (1988). Foundations and grand challenges of artificial intelligence: AAAI presidential address. *AI Magazine*, 9, 9–21.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258–276.

- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.
- Savage, L. J. (1972). *The foundations of statistics*. New York: Dover.
- Schmidt, F. L. (1971). The relative efficiency of regression and simple unit weighting predictor weights in applied differential psychology. *Educational and Psychological Measurement*, *31*, 699–714.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*, 610–628.
- Selten, R. (2001). What is bounded rationality? In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 13–36). Cambridge, MA: MIT Press.
- Serwe, S., & Frings, C. (2006). Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making*, *19*, 321–332.
- Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, *81*, 75–86.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1–19.
- Smith, V. L. (2003). Constructivist and ecological rationality in economics. *American Economic Review*, *93*, 465–508.
- Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Clarendon.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, *69*, 213–225.
- Todd, P. M. (2001). Fast and frugal heuristics for environmentally bounded minds. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 51–70). Cambridge, MA: MIT Press.
- Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 287–308). New York: Oxford University Press.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299.
- Wübben, M., & Wangenheim, F. v. (2008). Instant customer base analysis: Managerial heuristics often “get it right”. *Journal of Marketing*, *72*, 82–93.
- Yee, M., Hauser, J., Orlin, J., & Dahan, E. (2007). Greedoid-based non-compensatory two-stage consideration-then-choice inference. *Marketing Science*, *26*, 532–549.