

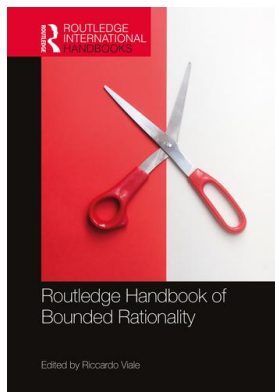
This article was downloaded by: 10.2.97.136

On: 22 Mar 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



Routledge Handbook of Bounded Rationality

Riccardo Viale

Philosophical naturalism and bounded rationality

Publication details

<https://test.routledgehandbooks.com/doi/10.4324/9781315658353-4>

Thomas Sturm

Published online on: 29 Oct 2020

How to cite :- Thomas Sturm. 29 Oct 2020, *Philosophical naturalism and bounded rationality from:* Routledge Handbook of Bounded Rationality Routledge

Accessed on: 22 Mar 2023

<https://test.routledgehandbooks.com/doi/10.4324/9781315658353-4>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

3

PHILOSOPHICAL NATURALISM AND BOUNDED RATIONALITY

Thomas Sturm

Introduction

Rationality (or reason¹) has often been declared to be a central topic of philosophy: as the source of principles either for metaphysics and ethics, or for epistemology and philosophy of science. Some authors even identify philosophy with the theory of rationality (Habermas, 1981, p. 16; Putnam, 1981, p. 113; Nozick, 1993, p. xi; Nida-Rümelin, 1996, p. 73; Grice, 2001, p. 4). However, most philosophers nowadays understand that it is not only philosophers who have developed our understanding of rationality. Beginning in the Enlightenment, and to an ever greater extent over the last century, sciences such as mathematics, statistics, economics or psychology have (partly) conquered this territory (Gigerenzer et al., 1989; Erickson et al., 2013; Sturm, forthcoming). The contributions on bounded rationality from Herbert Simon through to Gerd Gigerenzer, and their collaborators and followers, belong to this trend too.

The idea that theories of rationality obtain support from science should be grist to the mill of the philosophical doctrine of *naturalism*. While there are many versions of naturalism, they all share the view that the relation between philosophy and the sciences is (at least) one of close allies, and that the sciences contribute through their theories and methods to the solution of philosophical problems. Anti-naturalists, in turn, defend a separation of philosophy and the sciences, insisting that philosophy has *sui generis* methods for dealing with its own distinctive problems.

Both these opposing positions have been applied to rationality (Sober, 1981; Putnam, 1982; Stich, 1985, 1990, 1993; Hauptli, 1995; Pacho, 1995; Stein, 1996; Chiappe & Verwaeke, 1997; Bermúdez & Millar, 2002). Meanwhile, very few philosophers have drawn on Simon's (1956, 1957) work on *bounded* rationality (Giere, 1985, 1988; Cherniak, 1986; Gigerenzer & Sturm, 2012; deLanghe, 2013; Hahn, 2013, Chapter 8).² This is surprising, since bounded rationality (1) is an empirically grounded account of reasoning; and (2) does not merely have descriptive or explanatory aspirations, but normative or prescriptive ones as well. Thus, this notion of bounded rationality offers a path towards a naturalistic account of rationality—one that addresses the important issue of the normativity of rationality, which is often seen as a major challenge to any such naturalization.

No comprehensive study of the philosophical aspects of bounded rationality exists. Such a study would, at least, have to (1) explicate and scrutinize the presuppositions of the concept

(such as its descriptive as well as normative aspirations, or its difference to non-bounded conceptions of rationality); (2) determine areas in philosophy where models of bounded rationality might be used, and assess such uses where they exist; and (3) address the fundamental debates over its theoretical and methodological adequacy in psychology, economics, sociology, and political science, as well as the potential and limits of its interdisciplinary uses. Such a study cannot be carried out here, I shall include selected aspects of (1) and (2). I begin by outlining three systems of reasoning that are constitutive of the so-called “standard picture” of rationality, and highlighting two major criticisms of that picture, both associated with bounded rationality. Then, I introduce some major assumptions of naturalism and the challenges they face. Finally, I discuss the prospects and limits of bounded rationality for naturalism in epistemology, with occasional considerations concerning the philosophy of science.

The “standard picture”: three normative systems of rationality

We try to reason well, whether to convince others or ourselves of a point of view, to discover new knowledge that allows us to defend our pet beliefs, or to make good decisions. Our reasoning certainly can be either good or bad. However, this can only be the case if there are standards or norms of reasoning that have been agreed upon. What are they? According to Stein’s (1996) useful term, the “standard picture” of rationality consists of a set of three normative systems of judgment and decision making that are dominant nowadays in numerous areas, and often opposed by defenders of bounded rationality.

One of those systems is derived from the revolution in logic that started in the nineteenth century with the work of the philosopher-mathematician Gottlob Frege (1848–1925). If we consider this, it will quickly lead us to the dominant view concerning the relation between reasoning and normativity. Frege’s work influenced subsequent work in logic, from that of Bertrand Russell and Alfred N. Whitehead through to that of Alfred Tarski and Alan Turing. Thus, Frege’s work even paved the way for modern computers—and for Simon and Newell’s “Logic Theorist” program (Newell et al., 1958). Frege’s main project was to show that a part of mathematics, namely, arithmetic, could be reduced to logical laws. This is the doctrine of *logicism*. Since the existing logics, as Frege found them, were insufficient for such a reduction, he radically overhauled them by creating an axiomatized predicate calculus with innovative tools for complex quantification. While his logicism was—as he himself later admitted—unsuccessful, the central achievements of his new logic are still alive and well today.

Frege was aware of views that drew no sharp distinction between logic and psychology. He rejected them by claiming that the *validity* of logical laws is independent of how people *actually* think and reason (Frege, 1966). This is his *anti-psychologism*, which has remained a forceful reason for anti-naturalism until today. Frege’s main basis for this stance was that actual reasoning is subject to mistakes and fallacies, and so cannot ensure what logic guarantees. Logical rules determine whether reasoning from premises to conclusions is truth-preserving. Whether the premises of an argument are true or not cannot be decided by logic alone; but logic tells us which patterns of inference guarantee that *if* the premises are true, *then* the conclusion will also be true. The validity of logical laws has to do, among other things, with the meaning of logical connectives such as ‘ \otimes ’ (interpreted as a specific version of the ordinary “if-then” conditional) and ‘ \oslash ’ (negation). Thus, *modus ponens* (1. $p\otimes q$; 2. p ; 3. Therefore, q .) or *modus tollens* (1. $p\otimes q$; 2. $\oslash q$; 3. Therefore, $\oslash p$) are deductively valid under any interpretation of the propositional variables p and q . Of course, humans sometimes violate such rules. For instance, there is the fallacy of “denying the antecedent”: 1. $p\otimes q$; 2. $\oslash p$; 3. Therefore, $\oslash q$. Consider for example:

If you are a banker, then you have a regular income. You are not a banker. Therefore, you do not have a regular income.

Clearly, you do not have to sell your soul to monetary businesses to have a regular income. In other inferences, the mistake is less easy to detect. Alan Turing considered the following argument: “If each man had a definite set of rules of conduct by which he regulated his life he would be no better than a machine. But there are no such rules, so men cannot be machines” (Turing, 1950, p. 452). Again, this is an example of the same fallacy. But maybe what those who propose such an argument really mean is:

Only if each man had a definite set of rules of conduct by which he regulated his life he would be no better than a machine. But there are no such rules, so men cannot be machines.

In this case, the inference would not embody “denying the antecedent”. But surely people can mix up *p*’s and *q*’s, neglect logical meaning, and thus end up committing a fallacy.

While logicism concerns the foundations of arithmetic, anti-psychologism concerns the basis of logic. Even psychologists such as Peter Wason (1966) and many others have accepted anti-psychologism by testing human reasoning against logical laws taken as unquestioned normative yardsticks. Neither did Simon and Newell question the validity and applicability of formal logic—the Logic Theorist was constructed to prove numerous theorems in Russell and Whitehead’s *Principia Mathematica* (the artificial intelligence thesis), and Simon and Newell also viewed the program as a model of human reasoning (the so-called information-processing thesis): the Logic Theorist “provides an explanation for the processes used by humans to solve problems in symbolic logic” (Newell et al., 1958, p. 163). Thus, the logical model was projected into the mind, not the other way around (Gigerenzer & Sturm, 2007, pp. 325–327).

Anti-psychologism is often described as being based on a rejection of the Is–Ought fallacy (noted in Hume, emphasized by Kant, used and discussed further in later debates on theoretical and practical rationality; see Sturm, forthcoming): one cannot infer normative judgments from descriptive ones. For instance, from the descriptive claim that people often engage in tax evasion, it does not follow that one should engage in tax fraud. From the claim that people often do not blame others for engaging in tax fraud, it does not follow that one should not be blamed for engaging in tax fraud. Frege’s anti-psychologism, however, is *not* derived from rejecting the Is–Ought fallacy. Contrary to what is sometimes asserted (e.g., Notturmo, 1985), Frege did not view the laws of logic as inherently normative, but as descriptive. However, descriptive of *what*? He claimed that logical rules or patterns of inference are universally and timelessly valid. His objection to psychologism, then, addressed a specific version of this doctrine: attempts to explain logical “laws of thought” as laws of actual thinking, with the latter being understood as a chain of “subjective” mental representations. Against this version of psychologism, Frege argued that such representations are too variable to constitute a possible basis of logical laws, given their universality and timeless validity. That is why, in order to justify such laws, Frege (1966, 43) required that we assume a “third realm” (*drittes Reich*) beyond the material world and our subjective mental representations. This argument is disputed, but it cannot be discussed here (cf. e.g., Sober, 1978; Carl, 1994, Chapter 2).³ What matters here is a straightforward point: since we can make mistakes in our reasoning, we ought to correct them, and that may involve logical laws. To deny this force of logic by claiming that our actual reasoning sometimes works in different ways is to fall prey to various kinds of psychologism.

Still, we will see that under certain conditions, the charge of psychologism can be avoided, and assumptions of bounded rationality play a role in this.

Anti-psychologism became popular beyond logic, especially through the epistemology and philosophy of science of the Vienna Circle and Popper's Falsificationism (Peckhaus, 2006). However, not all reasoning, whether ordinary or scientific, is deductive. Many non-deductive inferences, for instance, about scientific hypotheses, use probabilities. This leads to a second normative system that is important for the standard picture: theories of probability as developed by, e.g., Rudolf Carnap or Hans Reichenbach, alongside with mathematicians and statisticians.⁴ Their theories tried to do for "inductive" logic what Frege had achieved for deductive logic: an axiomatic system with clear tasks, limits, and structure, only one that normatively guides probable reasoning. While the mathematics of probability was axiomatized by the mathematician Andrej Kolmogorov (1903–1987), the meaning of probability remains disputed, and thereby so do its proper applications. There are, for instance, logical theories aimed at determining degrees of confirmation above 0 for scientific hypotheses, relative to given empirical evidence (Carnap); the so-called frequentism, which views probabilities as the limiting frequency of outcomes in long-run or potentially infinite series of similar events (Reichenbach); or the propensity theory, which takes probabilities to be properties that are inherent to sets of repeatable conditions (Karl Popper). However, none of these theories is without its problems; and none of them can claim to represent "the" meaning of 'probability': a term which is ambiguous and can serve different functions (Gigerenzer et al., 1989; Gillies, 2000). A fundamental pluralism with regard to probability is widespread today. What is more, while Carnap and others assumed that Fregean anti-psychologism concerning logic could be expanded into epistemology or philosophy of science, today it is far from clear that such a move comes without costs (Sober, 1978; Peckhaus, 2006). It would require all philosophical claims about knowledge or science to be formal or logical truths, which is highly implausible.

Third, and finally, since the mid-twentieth century, theories of rational choice have emerged, beginning with John von Neumann and Oscar Morgenstern's landmark *Theory of Games and Economic Behavior* (1944), which was itself influenced by the rise of logic. Moreover, rational choice owes much to economic theories of utility maximization that emerged in previous centuries. It seems plausible to say that if you are a rational agent, you ought to maximize your (expected) advantage, given what you believe the probability of certain outcomes to be.

So, these normative systems coalesced into the standard picture of rationality: if you are perfectly rational, you have to reason in line with logic, probability, and decision theory. Even today, these formal and optimizing theories are viewed even by many as (1) a descriptive model for human reasoning and, perhaps more importantly, (2) the normative yardsticks for such reasoning.⁵ However, questions concerning the normative validity of the standard picture arise: What justifies the rules of each of these theories? And why should we reason in line with them in particular situations?

Two objections to the standard picture

I now consider two influential objections to the standard picture, both related to bounded rationality. As the term "bounded rationality" is often meant to include *all* empirical theories that try to account for informal and non-optimal reasoning using both lazy and hard constraints (Oaksford & Chater, 1991), the objections differ considerably. This, in turn, leads to different connections that one can draw between bounded rationality and naturalism.

Objection 1: Insofar as the formal theories were interpreted descriptively—not in the Fregean sense of describing a realm of logical entities and their formal relations, but as being descriptive

of human thinking and decision-making—one could study *empirical* counterexamples to them. This led to the demand to build an alternative account of judgment and decision-making. Thus, Wason's (1966) four-card test aimed to refute Piaget's theory of cognitive development, according to which we master propositional calculus from age 12 onwards. Similarly, the program of "heuristics and biases" (henceforth, HB), developed by Daniel Kahneman and Amos Tversky, attempted to show that the view of "man as an intuitive statistician" (Peterson & Beach, 1967) whose behavior approximates the canons of probability was mistaken. Reasoning fallacies and mistakes that the HB program detected include conjunction errors (the "Linda problem"), neglecting the obligation to use sufficiently large statistical samples ("law of small numbers"), the base rate fallacy, and so on (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1983).

Kahneman claims that Simon's work is in line with the HB approach:

[Tversky and I] explored the psychology of intuitive beliefs and choices and examined their bounded rationality. Herbert A. Simon ... had proposed much earlier that decision makers should be viewed as boundedly rational, and had offered a model in which utility maximization was replaced by satisficing. Our research attempted to obtain a map of bounded rationality, by exploring the systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models.

*Kahneman, 2003, 1449; see also Kahneman, Slovic,
& Tversky, 1982, pp. xi–xii*

Kahneman's view that humans are often irrational results from adopting rules of the standard picture as yardsticks. Our intuitive judgments and decisions then have to be explained in terms of so-called heuristics (e.g., "representativeness" or "availability"). Understood this way, "bounded rationality" implies that standard "rational models are psychologically unrealistic", and "maps of bounded rationality" are supposed to account for deviations from the "rational agent model" (Kahneman, 2003, p. 1449).

Objection 2: Formal and optimizing rules set substantive norms for ideal reasoners, i.e. reasoners equipped with infinite time, memory, and other resources. But, Simon said, we must distinguish between substantive and "procedural" rationality (a distinction that paved the way towards bounded rationality; see Erickson et al., 2013, Chapter 2). It is one thing to judge whether an outcome is reasonable in the light of a standard norm; it is quite another to judge whether an outcome is reasonable in the light of whether the process that leads to it is overly costly.

Importantly, there are stubborn problems of *computational intractability*. Cognitive processes which require resources that increase at an exponential rate (i.e., 2^n , or more) are regarded as creating such problems. Consider the traveling salesman problem, i.e. finding the optimal route between a number of locations when one travels to each location only once and the whole tour is as short as possible. Calculating the optimal route between e.g., 50 cities requires calculating no less than 300,000 possible routes. To date, no human, and no computer can do that. There is a similar situation for tasks such as finding the optimal path for winning in chess, or making investment choices, and other daily and scientific problems (Michalewicz & Fogel, 2000). Hard computational intractability may make it impossible to use the norms of the standard picture. Even if the formalistic and optimizing approach to rationality produced rules that were valid within the axiomatic systems to which they belong, we may have to give up the close connection between

the formal validity and the rationality of patterns of inference. There is, then, a gap between formal (including optimizing) rules and rationality.

Simon and others have seen this gap as a major reason for developing models of “bounded rationality”, in order to figure out how people actually do reason, but also how they can reason well:

Bounded rationality dispenses with the notion of optimization and, usually, with probabilities and utilities as well. It provides an alternative to current norms, not an account that accepts current norms and studies when humans deviate from these norms. Bounded rationality means rethinking the norms as well as studying the actual behavior of minds and institutions.

Gigerenzer & Selten, 2001, p. 6

One of Simon’s most influential ideas was to replace the goal of maximizing expected utility by “satisficing”. In Gigerenzer’s view, reasoning (often) uses fast and frugal heuristics (henceforth, FFH) that also work with little information and computation, such as the “recognition heuristic”, the “fluency heuristic”, or “take the best” (for an overview of ten well-studied heuristics, see Gigerenzer & Sturm, 2012, p. 249f.). That FFH do not merely describe behavior but should, in a number of cases, also guide our judgments and decisions is justified if and insofar as reasoning processes express a certain fitness function between mind and environment (Gigerenzer, Chapter 2 in this volume). That function needs to be discovered empirically.

Understood this way, bounded rationality questions the sharp divide between the descriptive and the normative—quite different from Kahneman’s understanding of the concept. However, no Is-Ought fallacy thereby needs to be implied. Instead, underlying this blurring of the boundary is a legitimate principle of normativity: *ought implies can* (also already recognized by Kant). The intractability of many reasoning tasks restricts the applicability of certain formal and optimizing norms, and thus makes it reasonable to look for different, more feasible norms.

The difference between these two objections leads to two disputes. The first, purely empirical objection has prompted discussion on the extent to which humans are rational, asked in the light of norms of the standard picture of rationality. The results range from “completely irrational” to “sometimes irrational”. In contrast, the second objection has raised the question of which norms should guide good reasoning in the face of widespread intractability and uncertainty: those of the standard picture, or those of bounded rationality understood normatively (i.e., Simon’s and Gigerenzer’s approach)? Or maybe both? As we will see next, philosophical projects of naturalizing rationality disagree here too.

Naturalism: its aims, scope, assumptions, and problems

Although it is considerably older, naturalism has become popular since the 1960s (Quine, 1969; Kitcher, 1992; Kornblith, 1994; Rosenberg, 1996).⁶ Naturalists object to “armchair” methods such as conceptual analysis, thought experiments, or appeals to intuition (DePaul & Ramsey, 1998; Kornblith, 2007), claiming that science provides better resources with its methods or theories, and viewing philosophy that isolates itself as sterile and obsolete. Some naturalists have even claimed that the traditional tasks and methods of philosophy will sooner or later be *replaced* by science (e.g., Quine, 1969).

Naturalism is a *metaphilosophical* thesis: a view about philosophy’s questions, methods, and aims. To be convincing, naturalism has to prove itself by being specific as to its domains, as to what claims it makes on what grounds, and as to how it defends itself against objections. Thus,

naturalism has been developed in metaphysics (Ladyman, 2007), philosophy of mind (Dennett, 1991; Dretske, 1996), ethics (Harman, 1977, Lenman, 2009), epistemology (Goldman 1986; Kornblith, 1993, 2002), philosophy of science (Giere, 1985, 1988, 1992; Laudan, 1987, 1990; Kitcher, 1993; Mi & Chen, 2007; Viale, 2013), and other domains as well (De Caro & MacArthur, 2004, 2010; Galparsoro & Cordero, 2013). Cutting across these divisions of domains, naturalists make different general claims. Two of their basic convictions are:

(ON) Reality consists only of natural objects, their properties, and relations: there are no unexplainable or irreducible supernatural entities or powers.

(MN) Philosophical questions can and should be answered by relying on the methods of the empirical sciences: there is no *a priori* philosophical knowledge and there are no special philosophical methods.

“ON” stands for “ontological naturalism”, and “MN” for “methodological naturalism” (De Caro & MacArthur, 2010; Galparsoro & Cordero, 2013). Naturalists usually have additional commitments, such as to Darwinism (Rosenberg, 1996) or physicalism (Papineau, 2015). Without them, (ON) might be so uncontroversial that most non-naturalistic philosophers would accept it too. We do not know what (ON) comes down to unless we have a specific scientific theory that explains the relevant objects and their properties and relations. If one defends (ON) about the mind, one could use theories rooted in biology or neuroscience. In the area that concerns us here, rationality, naturalists usually look to psychology, but also to other sciences, such as those just mentioned. Such an expansion also helps to distinguish between psychologism and naturalism. In any case, it all depends on the quality or validity of the specific theory the naturalist offers. Naturalists acknowledge the necessity of this task of specification, and many have invested much effort in connecting scientific theories to philosophical problems.

(MN) is controversial too. To begin, can *all* philosophical questions be answered by scientific methods? Or only *some*? If so, which ones? And can the sciences deliver? Questions concerning the nature of time, free will, knowledge, or the normative basis of morality certainly do not seem to be suited to empirical methods. With few exceptions (despite the recent development of “experimental philosophy”), philosophers do not perform observations, experiments, statistical analysis, and so on. As Nozick (1993, p. xi) has noted, what philosophers really love is not wisdom but *reasoning* as such: thinking about it as well as practicing it.

As if all that were not enough, naturalists have to face straightforward anti-naturalistic arguments, e.g., the claims that we have *a priori* knowledge that cannot be explained naturalistically (Putnam, 1982), that naturalism falls prey to naturalistic fallacies (Kim, 1988), that it leads to circularity or self-destructiveness (BonJour, 1994), triviality (Stroud, 1996), and a lack of clarity (Sklar, 2010). We will see that similar points come up when we look at naturalism about rationality. In particular, a central challenge naturalists have to face in our case should be clear: to provide a scientific account that treats the normativity of rationality as something within the natural world, and explainable by scientific theory. This is not an easy task.

Naturalism about (bounded) rationality

Naturalists have drawn a variety of connections between rationality and the sciences. They appeal to the HB approach in order to address problems of epistemic rationality (Goldman, 1986; Stich, 1990; Kornblith, 1993; Bishop and Trout, 2005) or decision making concerning scientific theories (Solomon, 1992; Kitcher, 1993). They deal with claims of evolutionary psychology to explain the nature and foundation of rationality (Stein, 1996); and occasionally

research on satisficing and FFH has been exploited as well (Giere, 1985, 1988; Cherniak, 1986; Gigerenzer & Sturm, 2012; Sturm, 2019; Hey, 2016).

Before we can move on, some clarification is needed. If naturalism about rationality merely required *some* sciences to be used in the detection, explication, and justification of the norms of rationality, then the standard picture might already count as naturalistic, since it involves results of mathematicians or economists. However, that would contradict (MN). The normative systems of the standard picture are all *a priori*, not the result of observation or experiment.⁷ One might, of course, suggest dropping the reference to empirical science, thus replacing (MN) by:

(MN*) Philosophical questions can and should be answered by relying on the methods and results of the sciences.

Stated differently, “naturalism” would just mean “scientism”. This move, however, trivializes naturalism about rationality. Since the non-empirical, *a priori* character of logical rules has always been a central point of conflict between psychologism and anti-psychologism, naturalism and anti-naturalism, we should not take it for granted that a formal-rules account of rationality is naturalistic enough. The same line of reasoning works, *mutatis mutandis*, for theories of probability and decision making. The fact that Kolmogorov was a mathematician and not a philosopher does not mean that his work provided a naturalistic account of probability theory.

(MN*) might, furthermore, undermine (ON): If naturalists simply accept what comes from the formal sciences, without giving a naturalistic explanation of the relevant abstract entities and rules, then they might be asked how they deal with Frege’s talk of a “third realm” of logical laws. (ON) must do better. An interesting naturalism about rationality should therefore build on genuinely *empirical* theories. No surprise, then, that many naturalists about rationality look to evolutionary and cognitive psychology, or to biology and neuroscience. For reasons of space, I cannot further discuss (ON) about rationality in this chapter (cf. Sober, 1981; Chiappe & Verwaeke, 1997).

Now, as seen above, the term “bounded rationality” has been used for both the HB program, and the Simon-Gigerenzer program of satisficing and FFH. Let us now consider each of these in turn.

Once the HB program had become popular in science, it was adopted by naturalists too. Many viewed it as the scientific version of the mundane view that humans are prone to errors and fallacies. What is more, some followed the interpretation according to which HB studies showed that humans are deeply irrational (e.g., Stich, 1985). Unfortunately, such naturalists had to recognize that results of the HB approach became increasingly contested. The criticism not only came from philosophical armchair arguments (though these could be influential: Cohen, 1981). Psychologists too objected that the data of HB studies were not stable enough to claim that human beings violate the standard picture of rationality, that the theoretical concepts (such as “representativeness” or “availability”) were not precise enough to permit interesting, testable predictions, or that the normative assumptions used in the experiments could be questioned as well (Gigerenzer, 1991, 1996; Lopes, 1991; Cosmides and Tooby, 1996). This psychological debate is called the “rationality wars” (Samuels, Stich, & Bishop, 2002; Sturm, 2012a).

Thus, even naturalists have to make hard choices: Do the empirical results really warrant the conclusion that humans are highly irrational? What methods, concepts, and normative assumptions should psychologists use? What do we even *mean* when we call judgments, or processes, rational or irrational? Hilary Kornblith (1993, Chapter 5), another naturalist, has

discussed such questions for the “law of small numbers” (Kahneman, Slovic, & Tversky, 1982, Chapter 2). He argues that the alleged irrationality of human beings is not supported by the relevant studies, and that we can be more optimistic if we apply a background assumption from evolutionary theory:

Just as our perceptual mechanisms are well adapted to the environment in which they typically operate and build in presuppositions about the environment which are typically true, so our inferential mechanisms may also be built around presuppositions about standard environments which allow us to gain information about those environments quickly and accurately.

Kornblith, 1993, p. 86

He also immediately warns his reader: “There is of course, no a priori guarantee that this is the right perspective on human inference, and we should not try to force the data into such a mould” (Kornblith, 1993). Thus, naturalists do what they should abhor, given (MN): they reflect on science from a critical, normative point of view—as Kornblith says, “we *should not try* to force the data into such a mould” (emphasis added). It is, of course, the existence of conflicting approaches in cognitive psychology, or disputes over the interpretation of data that prompt Kornblith to take sides. Still, a strict and consistent naturalist should want to explain this critical, normative perspective itself again in purely naturalistic terms, by looking to e.g., a biological or psychological theory of such a critical, normative point of view. Kornblith does not do that, nor do other naturalists who engage critically with the empirical literature. Thus, while (MN★) trivializes naturalism about rationality, (MN) leads either into a regress or it implies the acceptance of an unexplained non-naturalistic perspective.

One reasonable response here is that naturalism should give up the aim of ultimately *replacing* philosophy by science, and should settle for *cooperation* between the two (cf. Feldman, 2001):

(CMN) Philosophical questions can and should be pursued in cooperation with the sciences (e.g., by using their methods and/or results).

(CMN) does not demand that all normative claims be explained in non-normative terms. Moreover, it is compatible with a certain normative use of HB results. Consider that epistemology has different traditional tasks (Stich, 1993): the definition of the concept of knowledge (as attempted in Plato’s *Theaetetus* or in much current analytic epistemology), the refutation of radical skepticism about knowledge (as expressed in Descartes’ first *Meditation*), or also genuinely normative tasks, such as justifying or correcting our beliefs, or providing methods for discovering new knowledge. Naturalists, when appealing to cognitive psychology, often hope to profit for the third, normative tasks.⁸

Goldman’s (1986, Chapters 13, 14) approach is an excellent example of this. Among other things, he shows that formal rules in the standard picture of rationality often do not map onto reasoning tasks in a straightforward manner. For instance, the logical law of non-contradiction, $\mathcal{O}(\mathcal{O}p \dot{\cup} p)$, states that a proposition and its negation cannot both be true. From contradictions, anything whatsoever can be derived. We should, it seems, therefore avoid inconsistencies. But then, what do you do when you discover inconsistencies in your belief set? Are you obliged to give up the *whole* set? Certainly not: we should separate the wheat from the chaff. However, that can prove to be difficult, especially in the complex belief systems of science, where several beliefs are all tenuous but part of a systematic theory. Similar examples can be found for other logical rules, such as deductive closure, or for rules of probability. There is, therefore, no simple

derivation of norms of epistemic rationality from formal rules of the standard picture. This is more support for the gap mentioned earlier on.

To bridge that gap, Goldman (1986; 2008) maintains that a distinctively philosophical or “analytic” epistemology has to determine the criteria or goals of our epistemic endeavors. He defends *reliabilism*: our belief-forming mechanisms need not always guarantee truth, but must safeguard a good ratio of true and false beliefs. We can learn from psychology which mental processes or mechanisms are reliable, and which are not. Heuristics are rules of thumb that often lead to true beliefs, though they also lead to errors and fallacies; that is just an epistemic risk we have to live with. For instance, HB studies provide insight into how *not* to reason: we can, and we should, learn to control the workings of heuristics so as to avoid biases (cf. Kahneman, Slovic, & Tversky, 1982, Chapter 30; and “dual systems” theory, Evans, Chapter 10 in this volume).⁹ In Goldman’s cooperative naturalism, cognitive psychology will not replace philosophical epistemology, but neither will the latter thrive without the former.

What should we think of such an approach? To begin with, Goldman, like other naturalists (e.g., Giere, 1985), thinks of epistemic rationality as a means–ends notion: given certain epistemic goals, we ought to use such-and-such a method. This *instrumentalism* about epistemic rationality helps to justify certain methods over others. Since means–ends relations have to be discovered empirically, this also adds support to the idea that norms and methods can be justified empirically. However, while Goldman asserts that the goals have to be set by “analytic” epistemology, and while he favors a reliabilism according to which we ought to select methods that lead to a good ratio of true and false beliefs, not all naturalists agree. Some argue that we should not care whether our beliefs are true, or that the goals we should care about are primarily pragmatic ones (Giere, 1988; Stich, 1990). Furthermore, it is far from clear that instrumentalism is acceptable at all when it comes to *epistemic* rationality: whether or not epistemic reasoning is good or bad might be entirely independent of what goals we pursue. At least some of our beliefs can be found to be justified, or be rationally warranted, no matter what our specific epistemic or other goals are (Siegel, 1989; Kelly, 2003).

Another point is that Goldman (1986) does not show that naturalists have no options besides HB. We will see how the Simon–Gigerenzer conception of bounded rationality provides an alternative.¹⁰ What is more, for HB to actually work, HB studies would have to be uncontested with respect to their claims concerning cognitive mechanisms. They simply are not. Kahneman and Tversky have long been challenged to produce process models that have real explanatory value or that allow for precise, testable predictions. The heuristics they cite, such as representativeness or availability, look more like redescriptions of the behavior they are supposed to explain (Gigerenzer, 1996).

However, let us assume for the sake of the argument that the HB program could do better at the explanatory and predictive tasks, and focus again on the normative issue. Goldman himself has pointed to the gap between formal rules and rationality. He, therefore, cannot—unlike Kahneman and Tversky—look back to the standard picture to provide unquestionable normative yardsticks for reasoning. For Kahneman and Tversky, but also for Goodman and many other psychologists and philosophers, heuristics are viewed as being justified by an *accuracy–effort trade-off*: we use them because looking for information and computation can be too costly; we trade a loss in accuracy for faster and more frugal cognition. The difference between the HB approach and Goldman’s reliabilism is that the former views heuristics as second best, whereas the latter acquiesces in saying that there is nothing better. For the HB approach, heuristics are good as rules of thumb, but ultimately our epistemic evaluations should be made in the light of standard, formal, and optimizing rules. For Goldman, we cannot but balance the costs and

benefits of reasoning mechanisms, and we should adopt an instrumentalist attitude towards the methods which we view as norms of epistemic rationality.¹¹

If one turns to the Simon–Gigerenzer line of thought, then Goldman’s line of argument can be disputed. Naturalists occasionally cite computational intractability (Goldman, 1986, p. 282), Simon’s satisficing (Giere, 1985; Cherniak, 1986; DeLanghe, 2013), and attempts to make use of FFH in scientific theory choice (Hey, 2016; for limits concerning this area, see Nickles, 2016). However, they do so in less systematic and comprehensive ways. Actually, around the time when Gigerenzer was developing his work on FFH and “ecological rationality” (in the early 1990s, that is), Kornblith was independently taking a similar step by suggesting, as cited above, that “our inferential mechanisms may also be built around presuppositions about standard environments which allow us to gain information about those environments quickly and accurately” (Kornblith, 1993, p. 86). FFH work well when they are used in the right environments; outside them, they may lead to errors. Consider the recognition heuristic (Goldstein & Gigerenzer, 2002). The surprising result is that subjects estimate, e.g., city sizes or winners of sports tournaments better when they know less. Thus, German subjects judge the size of US cities better than US subjects, and vice versa. Recognition by name alone works very well here. So, the heuristic works if and only if the environment—here, media and other information channels—guarantee that name recognition is systematically correlated with a correct estimation of the relevant criterion. Such a simple rule works better than trying to search for much more information. Similarly so, *mutatis mutandis*, for other FFH.

There should be nothing mechanical in the normative use of heuristics: we should not rely on them blindly. Consider one of the original recognition experiments, the comparison of the sizes of San Diego and San Antonio. When the study was carried out, San Diego was larger; by 2010, San Antonio had overtaken San Diego. That change in the environment was not yet reflected in German media channels. When German subjects can judge only by name recognition, they will probably still judge San Diego to be the larger city, and thus make a mistake. While FFH can be extremely useful for many reasoning tasks, an alert reasoner will use them with caution. We have to continue to think critically for ourselves.

At the same time, FFH can offer to naturalism—specifically, to (CMN)—more than a merely instrumentalist justification of norms: not only does the connection between means and (given) ends need to be studied empirically, but the relation of fitness between a heuristic and the environment in which we can, and should, legitimately use it, is a matter of empirical investigation too. At the same time, there are limits to the usefulness of FFH:

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.

Gigerenzer & Sturm, 2012, p. 244

This compatibility between the standard and the bounded conception of rationality can be deepened in two regards. First, one cannot even *formulate* FFH without basic concepts and rules of formal logic. The recognition heuristic requires a minimal grasp of the form of the “if – then” conditional: “If one of two objects is recognized *and* the other is not, *then* infer that the recognized object has the higher value with respect to the criterion.” Other rules, such as “take the best”, in addition, require the ability to master disjunction, and so on. Logical notions and rules are *built into* the very formulation of FFH. The gap between formal rules and rationality

notwithstanding, a minimal dependence of bounded rationality on basic logic is unavoidable (for a similar view, cf. Cherniak, 1986).

Second, consider the claim of Gigerenzer's program that heuristics are sometimes as good as, and sometimes even beat, probabilistic norms such as regression or Bayes' theorem. In accordance with what standard is such a claim being made? Clearly, each attempt to justify a heuristic as normatively valid requires some standard against which the heuristic is tested. More specifically, when defenders of bounded rationality claim that FFH outperform classical rules of the standard picture of rationality, the only way to prove this is by comparing how FFH fare in comparison with rules of probability or statistics. Accordingly, "take the best" is checked against actual frequencies, and other heuristics against, say, the benchmarks of Bayesian rules (Martignon & Blackmond Laskey, 1999; Martignon & Hoffrage, 1999). FFH can only be normatively valid if they *compete* successfully with such rules. Both these points mean that there is, at the normative level, a compatibility and even complementarity between the standard and the bounded conceptions of rationality (for more on this, see Sturm, 2019).

Conclusion: for a critical naturalism about rationality

Time to take stock. I have tried to show how productive a careful analysis of the perplexing historical, scientific, and philosophical interactions between the standard picture of rationality, the psychology of reasoning, models of bounded rationality, and philosophical naturalism can be. The standard picture arose through the coalescence of logic, probability, and rational choice theory through the twentieth century. That picture then was criticized by both the Kahneman–Tversky HB approach, which emphasizes that humans are lazy reasoners, and by Simon and later on by Gigerenzer and his collaborators for being too mindlessly formal and optimizing in terms of what the proper norms of good reasoning are. Philosophers who defend naturalism in epistemology and philosophy of science first quickly took up the HB approach, but did so in problematic ways that threatened their own naturalism.

Simon's satisficing and Gigerenzer's FFH are better allies of philosophical naturalism. Four points can be highlighted to summarize the results concerning this claim. First, a methodological naturalism about epistemic rationality can be based on this particular version of bounded rationality, given that means–ends as well as mind–environment relations have to be studied, both of which have to be determined empirically. Second, such naturalism about rationality and, in particular FFH, can be normative, but only if heuristics are used in mindful, deliberate ways, at least in principle. Third, such naturalism is limited in its scope and potential applications, and it should be fully reflexive with regard to these limits. It should live in coexistence with the standard account of rationality by leaving, for instance, computationally tractable tasks to the latter. It even requires the standard account to some extent, since logical notions and rules are required to formulate FFH, and theories of probability in order to assess their validity. Fourth, and finally, bounded rationality cannot be expected to provide a basis for the naturalists' ontological claim (ON) concerning rationality. Research into bounded rationality is not yet mature enough to deliver a systematic theory that could be used to explain the nature of rationality in general. Maybe it never will be.

Given these four points, we can justifiably call the ensuing position, while tipping our hats to Kant, "critical naturalism". It is naturalistic because of the first point, namely, its empirical methodology, and critical in the Kantian sense due to the three other points: We should use heuristics in reflective and limited ways, and we should avoid rushing towards ontological claims that are not sufficiently warranted. But even if one does not share this refined version of naturalism, what matters is that epistemologists and philosophers of science have, up to now,

all too rarely used bounded rationality as understood by Simon and by Gigerenzer and his colleagues. More work can and should be done in this area.

Acknowledgments

For comments and discussions, the author is grateful to Riccardo Viale and Gerd Gigerenzer. Christopher Evans helped to improve the language of this chapter, and gave several highly useful recommendations concerning content too. This work was supported by the Spanish Ministry for the Economy, Industry and Competitiveness (MINECO) through the research project “Naturalism and the sciences of rationality: an integrated philosophy and history” (FFI2016–79923–P).

Notes

- 1 I cannot here consider the view that these terms are not synonymous, but see Sturm (2018).
- 2 One might think that, given its title, Stich (1993) belongs into this camp too. However, Stich only uses Simon’s work in artificial intelligence (AI) to explain scientific discovery (Langley et al., 1987). AI and bounded rationality were Simon’s two major research agendas; but they had no internal connection that would necessitate dealing with the former here as well. Therefore, I will leave aside Simon’s research in AI, and Stich’s use of it to naturalize epistemology.
- 3 Frege’s distinction between the material realm, the realm of subjective mental representations, and the realm of objective thoughts bears similarities with Karl Popper’s equally notorious division into world 1 (the physical world), world 2 (the mental world), and world 3 (the world of objective knowledge) (Popper, 1972). However, Popper’s division is historically probably more influenced by his teacher, the psychologist Karl Bühler and his theory of basic functions of language (Sturm, 2012b). Moreover, while Popper’s theory is clearly intended to be an ontological distinction, it is a disputed question whether Frege’s claim about the “third realm” was truly intended to be an ontological claim (for a less demanding, epistemological interpretation of Frege’s view, see Carl 1994, Chapters 2–4).
- 4 Of course, theories of probability were far longer in the making, famously starting with the Pascal–Fermat exchanges on games of chance in 1654 (Gigerenzer et al., 1989; Gillies, 2000).
- 5 To think that even today some experts view (a) to be true may seem false, given the heuristics-and-biases approach that is explained above. However, things are not so simple. In economics, whether (a) is taken to be true or not depends a lot on whether standard norms, e.g., for rational choices, are viewed as descriptive of individual decisions (where it seems implausible), of aggregates of decisions of many individuals, or of economic structures. On this debate, see Ross (2014).
- 6 There is no reliable survey showing whether or not naturalism dominates current philosophy. A preliminary attempt was made by Bourget and Chalmers (2013). When asked about their metaphilosophical position, participants (mostly professional philosophers, mostly though not exclusively from Anglo-Saxon universities) chose naturalism first (49.8 percent), followed by non-naturalism (25.9 percent) and “other” (24.3 percent). Somewhat surprisingly, the same participants overwhelmingly denied essential claims of typical naturalistic positions, such as the rejection of *a priori* knowledge (71.1 percent said such knowledge exists), or the rejection of the distinction between analytic and synthetic propositions. Perhaps this reveals that the questions were problematic, but it is also possible that many philosophers hold unusual views (to say the least).
- 7 This may be questioned, e.g., in the light of historical considerations. Gigerenzer et al. (1989) argue that theories of probability were sometimes revised when they clashed with what educated minds thought about probabilities. Consider the famous St. Petersburg paradox: recognizing that no reasonable person was willing to invest infinite sums of money in a coin tossing game where the expected gains were infinite did not lead to the judgment that such a person was mistaken. Instead, it led Daniel Bernoulli to substitute the concept of expectation with that of monetary values, “with the awareness that the richer you are, the more it takes to make you happy” (Gigerenzer et al., 1989, p. 15). However, it would be questionable to claim that Bernoulli had attempted to justify a normative rule by means of observation or experiment. One might see this, rather, as an instance of applying the method of *reflective equilibrium*: we develop our normative theories by considering whether instances of good reasoning fit with our best principles, and vice versa; and we adjust them when conflicts arise (see also Cohen, 1981).

- 8 They also try, of course, to profit from psychology for other tasks of epistemology, for instance, in providing a descriptive account of knowledge or cognition. This is what Quine's (1969) original proposal for a naturalistic epistemology comes down to, or what "evolutionary epistemologists" have developed. However, such *descriptive* naturalistic epistemology faces the challenge that it simply changes the game by not taking seriously enough the normative tasks of epistemology (Kim, 1988). Accordingly, in order to avoid being charged with changing the topic, a number of naturalists accepted the normative tasks too.
- 9 Again, this undermines excessive conclusions (e.g., Stich, 1985) on the basis of HB studies according to which humans are fundamentally irrational, a claim that is not entirely absent from Kahneman and Tversky either; see Sturm (2012a).
- 10 Goodman (2008) offers a more mixed assessment of the competing approaches.
- 11 In philosophy of science, Solomon (1992) has used the HB program to explain e.g., the geological revolution of the nineteenth century. She finds that Stich (1985) has argued convincingly against criticisms of the HB program, a judgment with which I disagree, though I cannot show this here. She also uses HB studies to explain the normative successes of choices made during the geological revolution, claiming that Alfred Wegner's choice of continental drift over contractionism was driven by aspects of representativeness (Solomon, 1992, p. 447), and that belief perseverance or availability played a role too. However, the HB program of testing how far people follow formal or optimizing rules makes sense only where such rules can be clearly stated. Only then can judgments of bias be justifiably made, and explanations in terms of heuristics be given. Suitable norms just do not exist for cases such as those Solomon discusses. Nor does Solomon explain why geologists came to agree so quickly that continental drift was the more adequate theory in the 1960s. As Solomon admits, her explanation is only tentative and perhaps not the whole story anyhow.

References

- Bermúdez, J. L., & Millar, A. (Eds.) (2002). *Reason and nature*. Oxford: Oxford University Press.
- Bishop, M., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment*. New York: Oxford University Press.
- BonJour, L. (1994). Against naturalized epistemology. *Midwest Studies in Philosophy*, 19, 283–300.
- Bourget, D., & Chalmers, D. (2013). What do philosophers believe? Available at: <https://philpapers.org/archive/BOUWDP>
- Carl, W. (1994). *Frege's theory of sense and reference*. Cambridge: Cambridge UP.
- Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.
- Chiappe, D. L., & Vervaeke, J. (1997). Fodor, Cherniak, and the naturalization of rationality. *Theory & Psychology*, 7, 799–821.
- Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317–331 (comments and responses, 331–359).
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? *Cognition*, 58, 1–73.
- De Caro, M., & MacArthur, D. (Eds.) (2004). *Naturalism in question*. Cambridge, MA: Harvard University Press.
- De Caro, M., & MacArthur, D. (Eds.) (2010). *Naturalism and normativity*. New York: Columbia University Press.
- deLanghe, R. (2013). Satisficing as an account of Kuhnian rationality. *Philosophy Study*, 3, 398–411.
- DePaul, M., & Ramsey, W. (Eds.) (1998). Rethinking intuition: The psychology of intuition and its role in philosophical inquiry. Lanham, MD: Rowman & Littlefield.
- Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown & Co.
- Dretske, F. (1996). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Erickson, P., Klein, J., Daston, L., Lemov, R., Sturm, T., & Gordin, M. (2013). *How reason almost lost its mind*. Chicago: University of Chicago Press.
- Feldman, R. (2001). Naturalized epistemology. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Available at: <http://plato.stanford.edu/archives/sum2012/entries/epistemology-naturalized/>
- Frege, G. (1966). Der Gedanke (1918–19). In G. Frege, *Logische Untersuchungen*, ed. G. Patzig (pp. 30–53). Göttingen: Vandenhoeck & Ruprecht.
- Galparsoro, J. I., & Cordero, A. (Eds.) (2013). *Reflections on naturalism*. Rotterdam: Sense Publishers.
- Giere, R. (1985). Philosophy of science naturalized. *Philosophy of Science*, 52, 331–356.

- Giere, R. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.
- Giere, R. (Ed.) (1992). *Cognitive models of science*. Minneapolis, MN: University of Minnesota Press.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond heuristics and biases. *European Review of Social Psychology*, 2, 83–115.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky. *Psychological Review*, 103, 592–596.
- Gigerenzer, G., & Selten, R. (2001). Rethinking rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality* (pp. 1–12). Cambridge, MA: MIT Press.
- Gigerenzer, G., & Sturmfels, T. (2012). How (far) can rationality be naturalized? *Synthese*, 187, 243–268.
- Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). *The empire of chance*. Cambridge: Cambridge University Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (Eds.) (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gillies, D. (2000). *Probability*. London: Routledge.
- 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, Germany: Mentis.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75–90.
- Grice, P. (2001). *Aspects of reason*. Oxford: Oxford University Press.
- Habermas, J. (1981). *Theorie des kommunikativen Handelns*. 1. Frankfurt/M.: Suhrkamp.
- Hahn, S. (2013). *Rationalität*. Paderborn: Mentis.
- Harman, G. (1977). *The nature of morality*. New York: Oxford University Press.
- Hauptli, B. (1995). *The reasonableness of reason: Explaining rationality naturalistically*. Chicago: Open Court.
- Hey, S. (2016). Heuristics and meta-heuristics in scientific judgment. *British Journal for the Philosophy of Science*, 67, 471–495.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93, 1449–1475.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty*. New York: Cambridge University Press.
- Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, 66, 612–640.
- Kim, J. (1988). What is “naturalized epistemology”? *Philosophical Perspectives*, 2, 381–405.
- Kitcher, P. (1992). The naturalists return. *Philosophical Review*, 101, 53–114.
- Kitcher, P. (1993). *The advancement of science*. New York: Oxford University Press.
- Kornblith, H. (1993). *Inductive inference and its natural ground*. Cambridge, MA: MIT Press.
- Kornblith, H. (Ed.) (1994). *Naturalizing epistemology* (2nd ed). Cambridge, MA: MIT Press.
- Kornblith, H. (2002). *Knowledge and its place in nature*. Oxford: Oxford University Press.
- Kornblith, H. (2007). Naturalism and intuitions. *Grazer Philosophische Studien*, 72, 27–49.
- Ladyman, J. (2007). *Everything must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Langley, P., Simon, H., Bradshaw, G., & Zytkow, J. (1987). *Scientific discovery*. Cambridge, MA: MIT Press.
- Laudan, L. (1987). Progress or rationality? The prospects for normative naturalism. *American Philosophical Quarterly*, 24, 19–31.
- Laudan, L. (1990). Normative naturalism. *Philosophy of Science*, 57, 44–59.
- Lenman, J. (2009). Naturalism without tears. *Ratio*, 12, 1–18.
- Lopes, L., (1991). The rhetoric of irrationality. *Theory and Psychology*, 1, 65–82.
- Martignon, L., & Blackmond Laskey, K. (1999). *Bayesian benchmarks for fast and frugal heuristics*. Oxford: Oxford University Press.
- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 119–140). New York: Oxford University Press.
- Mele, A., & Rawlings, P. (Eds.) (2004). *The Oxford handbook of rationality*. Oxford: Oxford University Press.
- Mi, C., & Chen, R. (Eds.) (2007). *Naturalized epistemology and philosophy of science*. Amsterdam: Rodopi.
- Michalewicz, Z., & Fogel, D. (2000). *How to solve it: Modern heuristics*. New York: Springer.
- Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, 65, 151–166.

- Nickles, T. (2016). Fast and frugal heuristics appraisal at research frontiers. In E. Ippoliti, F. Sterpetti, & T. Nickles (Eds.), *Models and inferences in science* (pp. 31–54). Heidelberg: Springer.
- Nida-Rümelin, J. (1996). Zur Einheitlichkeit praktischer Rationalität. In K.-O. Apel & M. Kettner (Eds.), *Die eine Vernunft und die vielen Rationalitäten* (pp. 73–90). Frankfurt a/M: Suhrkamp.
- Notturmo, M. A. (1985). *Objectivity, rationality and the third realm. Justification and the grounds of psychologism: A study of Frege and Popper*. Dordrecht: Nijhoff.
- Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.
- Oaksford, M., & Chater, N. (1991). Bounded rationality in taking risks and drawing inferences. *Theory & Psychology*, 2, 225–230.
- Pacho, J. (1995). *¿Naturalizar la razón?* Madrid: Sigo XXI editores.
- Papineau, D. (2015). Naturalism. In E. N. Zalta (ed.), *The Stanford encyclopedia of philosophy*. Available at: <http://plato.stanford.edu/entries/naturalism/entries/naturalism/>
- Peckhaus, V. (2006). Psychologism and the distinction between discovery and justification. In J. Schickore & F. Steinle (Eds.), *Revisiting discovery and justification* (pp. 99–116). Berlin: Springer.
- Peterson, C., & Beach, L. (1967). Man as an intuitive statistician. *Psychological Bulletin*, 68, 29–46.
- Popper, K. (1972). *Objective knowledge: An evolutionary approach*. Oxford: Clarendon Press.
- Putnam, H. (1981). *Reason, truth, and history*. Cambridge: Cambridge University Press.
- 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, *Ontological relativity and other essays* (pp. 69–90). New York: Columbia University Press.
- Rosenberg, A. (1996). A field guide to recent species of naturalism. *British Journal for the Philosophy of Science*, 47, 1–29.
- Ross, D. (2014). *Philosophy of economics*. London: Palgrave Macmillan.
- 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: Oxford University Press.
- Siegel, H. (1989). Philosophy of science naturalized? Some problems with Giere's naturalism. *Studies in History and Philosophy of Science*, 20, 365–375.
- Simon, H. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Simon, H. (1957). *Models of man*. New York: Wiley.
- Sklar, L. (2010). I'd love to be a naturalist-if only I knew what naturalism was. *Philosophy of Science*, 77, 1121–1137.
- Sober, E. (1978). Psychologism. *Journal for the Theory of Social Behaviour*, 8, 165–192.
- Sober, E. (1981). The evolution of rationality. *Synthese*, 46, 95–120.
- Solomon, M. (1992). Scientific rationality and human reasoning. *Philosophy of Science*, 59, 439–455.
- Stein, E. (1996). *Without good reason*. Oxford: Clarendon.
- Stich, S. (1985). Could man be an irrational animal? Some notes on the epistemology of rationality. *Synthese*, 64, 115–135.
- Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.
- Stich, S. (1993). Naturalizing epistemology: Quine, Simon, and the prospects for pragmatism. *Royal Institute for Philosophy Supplement*, 34, 1–17.
- Stroud, B. (1996). The charm of naturalism. *Proceedings and Addresses of the American Philosophical Association*, 70, 43–55.
- Sturm, T. (2012a). The “rationality wars” in psychology: Where they are and where they could go. *Inquiry*, 55, 66–81.
- Sturm, T. (2012b). Bühler and Popper: Kantian therapies for the crisis in psychology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, 462–472.
- Sturm, T. (2018). Rationalität versus Vernunft? Über eine Unterscheidung bei John Rawls (und anderen). In D. Sölch (Ed.), *Philosophische Sprache zwischen Tradition und Innovation* (pp. 211–233). Frankfurt a.M: Lang.
- 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, 190–209.
- Sturm, T. (Forthcoming). Theories of rationality and the descriptive/normative divide: A historical approach. In M. Knauff & W. Spohn (Eds.), *Handbook on rationality*. Cambridge, MA: MIT Press.
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.

- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: Conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.
- Viale, R. (2013). *Methodological cognitivism*, vol. 2: *Cognition, science, and innovation*. Heidelberg: Springer.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Wason, P. (1966). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273–281.