



Inquiry

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/sinq20>

The “Rationality Wars” in Psychology: Where They Are and Where They Could Go

Thomas Sturm^a

^a Universitat Autònoma de Barcelona, Spain

Available online: 18 Jan 2012

To cite this article: Thomas Sturm (2012): The “Rationality Wars” in Psychology: Where They Are and Where They Could Go, *Inquiry*, 55:1, 66-81

To link to this article: <http://dx.doi.org/10.1080/0020174X.2012.643628>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, 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 accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any 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.

The “Rationality Wars” in Psychology: Where They Are and Where They Could Go

THOMAS STURM

Universitat Autònoma de Barcelona, Spain

(Received 25 September 2011)

ABSTRACT *Current psychology of human reasoning is divided into several different approaches. For instance, there is a major dispute over the question whether human beings are able to apply norms of the formal models of rationality such as rules of logic, or probability and decision theory, correctly. While researchers following the “heuristics and biases” approach argue that we deviate systematically from these norms, and so are perhaps deeply irrational, defenders of the “bounded rationality” approach think not only that the evidence for this conclusion is problematic but also that we should not, at least not very often, use formal norms in reasoning. I argue that while the evidence for heuristics and biases is indeed questionable, the bounded rationality approach has its limits too. Most especially, we should not infer that formal norms play no role in a comprehensive theory of rationality. Instead, formal and bounded rules of reasoning might even be connected in a more comprehensive theory of rationality.*

Introduction

Current psychology of human reasoning constitutes a field where some of the most interesting debates about rationality are fought. In one such dispute, there are two main camps. On the one side, there is the “heuristics and biases” approach (“HB approach” for short), embodied most prominently in the work of Amos Tversky and Daniel Kahneman. This party claims, on empirical grounds, that human beings often and systematically violate norms of rationality that derive from formal logic, probability and decision theory. Instead, human judgment and decision making uses “heuristics”, rules

Correspondence Address: Thomas Sturm, Department of Philosophy, Universitat Autònoma de Barcelona, Edifici B, E-08193, Bellaterra (Barcelona), Spain. Email: tsturm@mpiwg-berlin.mpg.de

of thumb that work fairly well in some contexts but are not generally valid and thus lead to biases. On the other hand, an influential criticism of the HB approach—indeed, the only one which has prompted Kahneman and Tversky (1996) to a published reply—has been developed by defenders of theories of “bounded” rationality (“BR approach” for short). For this, the work by Gerd Gigerenzer and his colleagues stands out. They deny that human beings are systematically irrational in their judgment and decision-making. Also, they argue that it is wrong to take rules of logic, probability theory and statistics as unquestionable norms of rationality. Instead, we should construct norms out of an empirical study of the contents and contexts in which humans reason.

This debate has become so heated (see e.g., Kahneman & Tversky, 1996 and Gigerenzer, 1996a) that Samuels, Stich, and Bishop (2002) have called it the “rationality wars”. They claim that the debate is largely due to rhetorical excesses on both sides. Once the core assumptions of both programs are soberly compared, it becomes clear that the dispute isn’t really about much. For instance, they claim that the defenders of the HB approach do not all maintain that humans are *completely* and *unavoidably* irrational in their judging and deciding about logical, probabilistic or statistical tasks. According to this reconstruction of the debate, the issue at stake is simply how rational human beings are: *How many* standard rules of good reasoning do we actually apply? *How far* are we able to train ourselves to become sound reasoners? However, as I shall argue, the debate is also about different, more fundamental questions: What is the basis of norms of rationality? Can we use empirical investigations of human reasoning to answer this question? These questions are hard to answer, yet also important.¹ Contrary to what Samuels, Stich and Bishop argue, I insist that the theories of rationality assumed in the two research programs are substantively different, especially in their normative aspects. But, in contrast to both parties in the psychological debate, I doubt that these aspects are incompatible. A comprehensive theory of rationality could and should be composed of substantively different elements. We cannot build a solid home from bricks alone; we need cement, wood, glass, steel and other materials as well. (I should mention that another paper, co-authored with Gerd Gigerenzer [Gigerenzer & Sturm, 2011], shows that in important respects I agree with the BR approach, especially in that many norms of reasoning can indeed be validated empirically and why that is so. In the present paper, in contrast, I spell out where I disagree with the approach.)

In Section I, I illustrate the “heuristics and biases” approach more closely. In Section II, I explain how defenders of “bounded” conceptions of rationality criticize the HB approach by objections that are largely convincing. In Section III, however, I point out three limitations of the BR approach, leading to the idea of connecting it more closely with more traditional normative conceptions of rationality.

I. Are we highly irrational? The “heuristics and biases” approach

How could one find out whether human beings judge and decide in accordance with standard norms of rationality, such as those from formal logic, or probability and decision theory? A massive amount of the psychological literature from the last decades (see Lopes, 1991; Evans & Over, 1996) has applied the following procedure: Pick a particular norm, to be used in a task of more or less ordinary reasoning, and see how many experimental subjects apply it correctly. If they do that to a very large extent, then—given that the sample is representative—it may be inferred that human beings are rational in judgment and decision making; if not, they are irrational. Here is a concrete example, to which countless others could be added: the conjunction rule and the “Linda problem”.

The conjunction rule of probability theory states that an event A can never be less probable than a conjunction of the (independent) events A and B : $\text{Prob}(A) \geq \text{Prob}(A \& B)$. A typical psychological test concerning whether humans follow the conjunction rule goes as follows:

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination, social justice, and also participated in antinuclear demonstrations.

Which statement is more probable?

(T) Linda is a bank teller.

(T&F) Linda is a bank teller and active in the feminist movement.

Results: About 85% of 142 subjects chose the answer “T&F” and thus violated the conjunction rule (Tversky & Kahneman, 1983, p. 299).

The test for the conjunction rule has been varied in many ways. Different cases of persons, events etc. were described and then connected to various questions. Also, even when subjects were being informed about the correct answer, the tendency to produce incorrect answers decreased very little. Kahneman & Tversky (1996) insist that the results are stable (see also Kahneman’s views in Mellers, Hertwig & Kahneman, 2001). They explain this “cognitive illusion” by saying that subjects judge this way because of the *representativeness* of certain features of the description of Linda: If someone is concerned with issues about discrimination, social justice, and nuclear weapons, many people think that these properties are likely to be representative for her being a feminist as well. This heuristic makes subjects ignore the conjunction rule.

It is similar with many other reasoning tasks: Rather than using formal rules of logic, probability theory, or statistics and rational choice theory, humans seem to judge and decide on the basis of “heuristics” (rules of thumb which often work) which lead to “biases” (systematically incorrect results). In well-known studies with the *Wason Selection Task*, it has been argued that

most of us do not apply the rules of material implication: When asked to check instances of $P \rightarrow Q$, subjects overlook the relevance of instances of $\neg Q$ as potential falsifiers. For instance, subjects are given four cards with letters on one side (e.g. “E” and “C”) and numbers on the other side (“3” and “4”). Two cards are presented with the letter side up, two with the number side up. Then subjects are asked to turn over only those cards which are necessary for figuring out if the following statement is true: “If there is an E on the one side, then there is a 4 on the other side.” Most subjects pick out only the E-cards or the E- and the 4-cards (vowels and odd numbers). Less than 10% of the test subjects choose the right solution: the E- and the 3-cards. This is, in any case, the right solution if one follows standard propositional logic, which views statements of the form “If P , then Q ” as material implications: these are wrong if and only if P is true and Q is wrong. But subjects overlook the importance of the falsifying 3-card ($=\neg Q$), and are said to be biased towards looking for confirmation (Wason, 1966). Such mistakes, so the story goes, are not limited to laypeople. Statistically tutored Harvard medical students and staff members (!), when tested, were unable to correctly estimate the probability of breast cancer given certain diagnostic evidence; they committed *base rate neglect* (Casscells, Schoenberger, & Grayboys, 1978). They seemed to use an “availability heuristic”, using certain directly available items of evidence instead of base rates. We are prone to commit the *gambler’s fallacy* (thinking that the probability of an independent event is somehow affected by a series of previous events). And so on and so on (e.g., Tversky & Kahneman, 1974; Kahneman, Slovic & Tversky, 1982; Nisbett & Ross, 1980; Gilovich, Griffin & Kahneman, 2002).

What are the broader implications of these results? Many researchers have asserted outright that they have “bleak implications for human rationality” (Nisbett & Borgia, 1975, p. 935), or that “the selection task reflects [a tendency towards irrationality in argument] to the extent that subjects get it wrong . . . It could be argued that irrationality rather than rationality is the norm” (Wason, 1983, p. 59). Or, in a similarly pessimistic vein: “One might draw rather cynical conclusions . . . Human reasoning is fundamentally flawed” (Reisberg, 1997, p. 469; see also Gilovich, 1991; Sutherland, 1992; Piatelli-Palmarini, 1994).² Samuels, Stich and Bishop (2002) claim that Kahneman and Tversky have not made such radical claims. However, it must be said that they did draw very similar inferences. With respect to the error called the “law of small numbers”—the tendency to draw conclusions from very small samples rather than heeding the rule that only large samples will be representative of the population from which they are drawn—they have written: “The true believer in the law of small numbers commits his multitude of sins against the logic of statistical inference in good faith . . . His intuitive expectations are governed by a consistent misperception of the world” (Tversky & Kahneman, 1971, p. 31). In another study on decision theory, they have questioned the “favored position” of the “assumption of rationality” in

economics, allegedly treated by economists as “a self-evident truth, a reasonable idealization, a tautology, a null hypothesis”. It was, they have claimed, an assumption impervious to empirical falsification:

the assumption of rationality is protected by a formidable set of defenses in the form of bolstering assumptions that restrict the significance of any observed violation of the model. In particular, it is commonly assumed that substantial violations of the standard model are (i) restricted to insignificant choice problems, (ii) quickly eliminated by learning, or (iii) irrelevant to economics because of the corrective function of market forces. (Tversky & Kahneman, 1986, p. S273)

II. Maybe we are not so irrational after all: experimental artifacts and normative prejudices

Let us now look at how defenders of the bounded rationality approach in psychology attack the HB approach.³ Two of the most important objections here are the following:

- (1) The results concerning the systematic deficiencies of human reasoning are often, if not always, due to experimental artifacts.
- (2) Rationality is bounded: Psychologists should not take traditional formal rules of logic, probability theory and rational decision theory as normatively unproblematic, as is done in the heuristics and biases tradition. Rather, norms of reasoning are not valid independently of the contents and contexts in which they are applied; and to figure out the fitness between formal norms and contents and contexts of their use is a matter of empirical research.⁴

In what follows, I shall argue that (1) is convincing, but that (2)—while it contains important insights—also throws out some babies with the bathwater.

That there are experimental artifacts in the studies by Kahneman, Tversky and others can again be shown by reference to the “Linda problem”. The crucial point here is that the language used in these tests is by no means innocent, and that the understanding of core terms in the task questions can be influenced by prior information. In the “Linda problem”, Kahneman and Tversky presuppose that the terms “probable” or “more probable than” and “and” are all that counts when we test reasoning abilities. Moreover, they assume that these terms have to be understood such that “and” is the logical “AND” (&), and that “probable” conforms to principles of mathematical probability theory.

However, ordinary subjects do not understand them in these ways, especially not within the context of the Linda problem (Fiedler, 1988; Hertwig & Gigerenzer, 1999). The description of Linda as a politically sensitive person, active in the antinuclear weapons movement, and so on, pushes subjects

in the direction of certain interpretations of the statements (T) and (T&F). Given the forced alternative of the test, i.e., that condition that subjects must pick either “T” or “T&F”, many of them (20 to 50%) seem to infer that, for instance, “T” means to exclude “F” (i.e., $T = T \& \neg F$). Other subjects (10 to 20%), again, seem to understand “T&F” to mean “ $T \rightarrow F$ ”.⁵ Isn’t this irrational? If you mean by “rational” the standard formal rules of logic, probability theory, statistics and decision theory, then yes. But, Gigerenzer argues, that is not the only notion of rationality we have.⁶ Another notion includes our ability to draw semantic⁷ inferences. In our society, there is something like a correlation between having certain political attitudes and making choices for jobs. Bank tellers are not, in their overwhelming majority, people who care all day long about social justice, equality between the sexes, and so on. However, given the description of Linda, it seems implausible to many subjects that Linda is *not* F; hence they pick T&F. Others, again, may understand “T&F” roughly as “Well, if Linda has become a bank teller, then she is still concerned about discrimination, social justice, etc., and so it is more probable than not that she is a feminist too”. They might want to avoid the—indeed dubitable—judgment that Linda cannot be a feminist anymore once she has become a bank teller. Not only the logical and mathematical meaning of the terms “probable” or “and” matter in this task. The description of Linda is relevant as well, and reasonably so.

One can thus charitably reinterpret the violations of the conjunction rule such that subjects apply a different rational rule, namely a conversational rule (Hertwig & Gigerenzer, 1999, with reference to Grice’s (1975) “maxim of relevance”). It is, therefore, not necessarily adequate to label their inference fallacious or irrational. (To indicate another example: Moral or altruistic goals can reasonably ground decision making, even when the choices are at odds with formal decision theory. This has been shown by violations of “Property Alpha” or the “independence of irrelevant alternatives”; see Sen, 1993; Gigerenzer, 1998.)

Moreover, the alleged fallacies *are* avoidable, contrary to Kahneman and Tversky’s claim. We must only represent them in more transparent and unambiguous ways. For instance, representing the “Linda problem” in terms of a frequentist (rather than a subjective) interpretation of probability improves performances dramatically, even for statistically untutored subjects. Take the following task:

(Same description of Linda as before.)

There are 100 people who fit the description above. How many of them are (a) bank tellers, (b) bank tellers and active feminists?

In the answers to this format, the conjunction fallacy has been reduced from about 85% to 20% and less (Fiedler, 1988; Hertwig & Gigerenzer, 1999).⁸ Similar impressive results were achieved by using frequency formats in

Bayesian tasks, to the effect that subjects no longer overlooked the relevance of base rates (Gigerenzer & Hoffrage, 1995).

The same points can be made, *mutatis mutandis*, with regard to other alleged “cognitive illusions”. First, we can reinterpret data which apparently reveal fallacious reasoning such that it becomes clear that subjects are perhaps applying a rule which differs from the rule tested, but which may nevertheless be rational. Secondly, by using more transparent representations of reasoning tasks, even logically or statistically untutored subjects achieve correct answers at a much higher rate.

As should be pointed out, both considerations indicate only possibilities: It is not shown in such studies that subjects actually use rules of logic, probability theory, statistics, or alternatives such as conversational rules, and so on. Likewise, concerning the cases of avoiding “cognitive illusions” through more transparent representation, Gigerenzer et al. have not argued that the mind is actually a frequentist. To figure out what is actually going on is a rather intricate matter. In any case, it is quite possible that we are relatively rational after all, being able to learn rules of logic, probability theory etc. Since *ought* implies *can*, and *can* appears to be possible for us humans here, there seems to be no argument against the normative validity of those standard rules.

III. How far is rationality bounded?

However, Gigerenzer and with him many others⁹ do not think such moderation is enough. According to them, there is a higher price to be paid for what has been given so far. We must give up what Gigerenzer describes as “utopian” dreams about rationality: it is extremely unlikely that our minds were made to conform to truly universal norms of reasoning, and that the rules of logic, probability theory, and decision theory are these norms. Instead, human rationality is “bounded”.

What exactly does this mean? While there are various ways in which the conception of bounded rationality is spelled out, it should be said first that it does not mean that our reasoning abilities are limited and weak. This would, after all, come down to the claims of the HB approach.¹⁰ In contrast to such views, Gigerenzer points to a metaphor used by Herbert Simon (who coined the very terminology of “bounded rationality”): “Human rational behaviour is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor” (Simon, 1990, p. 7). In other words, one cannot leave out either the reasoning ability or the environment and still hope to adequately explain specific reasoning processes. Moreover, this is also normatively important. Instead of viewing standard formal rules of logic etc. as normatively valid (and therefore as appropriate tools for the psychological investigation of human reasoning), empirical arguments—specifically, arguments concerning the “contents and contexts” in which a certain rule works and those in which it does not—can and should be used to assess the validity of rules of reasoning.¹¹ So, the

claim that rationality is bounded is sometimes meant descriptively and sometimes normatively; and it is sometimes meant as a statement about reasoning *processes* or *activities*, and sometimes about *rules* of reasoning. The philosophically most challenging version of the BR approach would obviously be to claim that rules of reasoning are bounded, and that this is not merely a descriptive but also a normative claim. This contention is also at the center of the following considerations, in which I shall argue for certain limitations of the BR approach.

I present three kinds of considerations. First, I point out that Gigerenzer’s statements about the content-and-context dependency come in different varieties, some of which are acceptable, but not strong enough to support the BR approach in its rejection of formal rules as norms of rationality. Second, I consider some empirical studies of reasoning in order to show that arguments undermining the idea of a content-and-context-independent validity of rules are sometimes questionable. Third, I argue that fast and frugal heuristics and formal rules of logic, probability theory etc. play different and interlocking roles in a comprehensive theory of rationality.

As to the first point, consider the following general statements by Gigerenzer:

- (1) “I argued that *psychological principles are indispensable for defining and evaluating what sound judgment is*. Axioms and rules from probability theory and logic are, by themselves, indeterminate.” (1998, p. 464; emphasis added)
- (2) “The point I wish to defend . . . is that formal axioms and rules cannot be *imposed* as universal yardsticks of rationality independent of social objectives, norms and values; they can, however, be *entailed* by certain social objectives, norms and values. Thus, I am not arguing against axioms and rules, only against their a priori imposition as context-independent yardsticks of rationality.” (1996b, p. 320; emphasis added)
- (3) “My thesis is that traditional axioms and rules are incomplete as behavioral norms in the sense that *their normative validity* depends on the social context of behavior, such as social objectives, values, and motivations.” (Gigerenzer 1996b, p. 319; emphasis added)

These passages are not identical in meaning. Statement (1) is the weakest and can be accepted: Rules of logic or probability theory do not themselves contain rules or criteria for their correct *application*. In fact, no rule can, since that would lead into an infinite regress, as already no one less than Kant has pointed out (1781/1787, pp. A132–34/B171–74). Likewise, it is not true that because rules of logic or probability theory can be proven within formal calculi, they are therefore to be viewed as norms of good reasoning, because it is open how formal rules have to be mapped onto concrete reasoning tasks (see Goldman, 2008; Grice, 2001). Whether a rule should be applied to a certain task depends not only on the rule itself; we also need criteria for deciding

which rule out of a number of possible candidates should be applied to a given task. This may well depend on the subject's cognitive access to rules, his understanding of the task, his resources, the significance the task possesses (or should possess) for him, and so on. For all these issues, psychological knowledge is relevant. Being blind about this, and trying to mechanically apply certain formal norms to certain reasoning tasks, *is* a real weakness of the heuristics and biases approach. Still, let us not confuse, in general, questions of validity with questions of application.¹² In other words, if the BR approach would come down merely to claim (1), it would be restricted to determining conditions or criteria that tell us which rule is to be applied to a given problem.

Claim (2) states that the justification of a logical or probabilistic rule may be given by reference to the fact that it is *entailed* by certain values or objectives. Consider also how formal rules can be entailed by convincing instances of reasoning, as in the frequentist versions of the Linda problem or in Bayesian tasks. However, if viewing rules as being entailed by certain goals or successful instances of reasoning is meant to show that these rules are justified, this is a strategy that has nothing to do with the idea of bounded rationality or with validating norms by means of relating them to contents and contexts of reasoning, or to the environment of reasoners. Rather, it is a strategy that takes convincing instances of reasoning, showing what general rules are embodied in these instances, and then perhaps reflecting on whether these rules work in other instances as well, revising them if necessary, and so on. This is the well-known reflective equilibrium strategy (see e.g. Cohen, 1981). This is a serious option, but it does not entail that the validity of the norms as universal or formal rules must somehow be given up in favor of a bounded conception of reasoning norms.

Thus, only statement (3) expresses the view that the normative validity of rules depends on certain contents and contexts of reasoning—a much stronger claim than (1) or (2), one which, as I argue next, is not supported by the way Gigerenzer and others have attacked questionable uses of formal rules in the HB approach's studies on the Linda problem or the Wason Selection Task.

Consider first that there is an inconsistency in Gigerenzer's views concerning the Linda problem. He claims both that (a) answering T&F is *not* "a violation of the major view of probability, the frequentist conception" (1991, p. 92) *and* that (b) a representation of the test in frequency terms "makes the 'conjunction fallacy' largely disappear" (*ibid.*, p. 96) or causes "more correct answers" (1996a, p. 594, cf. p. 595). It seems to me that (b) is correct, for the reasons given by Gigerenzer himself. However, then answering that "T&F" is more probable than "T" *is* a violation of the conjunction rule, which is inconsistent with (a). Representing tasks in frequency formats improves performance; but saying this only makes sense if the two different representations of the task—one in terms of subjective probability, one in frequentist terms—concern the *same* rule. Gigerenzer must admit something like

that as well. Even though performance can be drastically improved by using frequency formats, he points out that still as many as 20% of the subjects do *not* follow the conjunction rule even under the improved experimental design. Because the other 80% of responses are regarded as improvements or as correct answers, the 20% must be on the side of those who still err.

One might object to this that we should not assume the conjunction rule to be *the* one and only rule that matters in the Linda problem. The conversational “maxim of relevance” might be a reasonable alternative. But then what is at stake here is, once again, not the *validity* of the conjunction rule, or its formal, content-independent validity, but whether it is the rule to be *applied* in this task. Whether reasoners should use the rule in any given reasoning task may—and probably does—depend on the context and content of the task in question, but that does not show that the validity of the rule does too.

Another problem may be illustrated with regard to the treatment of the Wason Selection Task by Leda Cosmides and John Tooby, accepted by Gigerenzer and others as an important argument in favor of the conception of bounded rationality (Gigerenzer, 1991; Gigerenzer, 1998; Gigerenzer & Hug, 1992). Cosmides and Tooby’s crucial point is that while subjects appear to reason badly when checking material implications with descriptive or indicative contents they improve when it comes to implications with deontic contents such as “cheater detection” (Cosmides, 1989; Cosmides & Tooby, 1996). Take a conditional such as “If someone is under 18 years old, (s)he is required to drink coke (a non-alcoholic beverage)”, where the consequent is a deontic statement. Here, subjects performed much better than in tasks where P and Q in $P \rightarrow Q$ were replaced by entirely indicative contents. (Subjects did somewhat better when the indicative conditionals were concrete rather than abstract, but they did still much better with deontic conditionals.) In the deontic conditionals, the large majority of subjects easily detected the relevance of instances of $\neg Q$ —whisky drinkers, say—as potential violators of the norm.

What explains this accuracy in cheater detection? Put briefly, Cosmides and Tooby’s account goes as follows. At some point of in the days of hunting and gathering, social cooperation and exchange became advantageous and even a distinctive step of the evolution of *homo sapiens*. Only few species have developed the ability to cooperate in ways that are expressive of genuine sociability. Cosmides and Tooby support this claim with reference to the work of Trivers (1971) and others (e.g., Axelrod & Hamilton, 1981; Axelrod, 1984; Maynard Smith, 1982) concerning evolutionary explanations of reciprocal altruism and cooperation. Such behavior can only constitute evolutionary stable strategies if the organisms possess the ability to spot those who cheat. Beings without such a capacity would be open to exploitation, and be selected out. This selection pressure, according to Cosmides and Tooby, has led to the evolution of a cheater detection module. It is successful in the area of such conditionals, with statements with these contents or purposes. There was no selection

pressure to understand all conditional statements properly, let alone to understand them along the lines of the material implication. This is why we are such fools in the original Wason Selection Task. According to these considerations, good reasoning depends upon the content and context of the task; in this case, upon specific *goals* we pursue due to our evolutionary adaptation.

However, this argument is problematic. It is an open matter whether subjects succeed so often because they have a specific cheater-detection ability. Fodor (2000) has argued persuasively that deontic conditionals are not sufficiently similar to indicative ones to make the case. In the latter case, what is asserted, and what has to be checked, is $P \rightarrow Q$ as a whole. By contrast, in the former case the requirement is contained merely in the consequent of the conditional, Q (the drinking part). The antecedent, P , merely specifies to whom the requirement applies (those under 18 years). This is a matter of fact about which there can be no sensible requirement. Subjects must, but also easily can, spot non-coke drinkers as potential violators. For this, they merely need to apply the Law of Non-contradiction, $\neg(Q \ \& \ \neg Q)$, which is much easier than to use the Law of Contraposition, $((P \rightarrow Q) \ \& \ \neg Q) \rightarrow \neg P$ needed in the case of indicative conditionals. Hence, the ability to reason successfully in cheater detection isn't necessarily due to the content (the goal), but might be explained by reference to a *formal* feature of the reasoning.

This consideration leaves open whether or not the apparently poor reasoning with indicative conditionals is irrational or not. Like in the Linda problem, perhaps subjects employ a different rule which may be reasonable as well. It has been proposed that, in the Wason Selection Task, subjects might understand the test sentence such that it does not refer to the four cards only, but as an empirical generalization: "For all x , if $Fx \rightarrow Gx$ ". For instance, they might consider checking whether all swans are white, and do so by looking at swans and white things, leaving out thereby all non-white things. What do you think needs more time, checking all the white things or all non-white things? (Oaksford & Chater, 1994; Botterill & Carruthers, 1999, p. 125) If that is so, then subjects might be viewed as in some sense rational. The difference to the Linda case is that here, the material conditional might still be the rule to be applied when it comes to the ultimate *justification* of the statement (as opposed to the pursuit of a hopefully less costly way for the selection of data). But even though subjects are, in this case, not perfectly rational, that by itself does not undermine the universal validity of the material conditional.

One might object at this point that the material conditional is a rather questionable thing. It is well-known that the rules for material implication run into paradox. A statement such as "If Rome is the capital of Mongolia, then the FC Barcelona will win the Champions League seven times in a row" is true simply because Rome isn't the capital of Mongolia. But the validity of the rules of material implication remains a matter for logicians, not for psychologists. The interesting question—and here Gigerenzer's opposition to a blind use of formal rules within psychology is quite in place—is why psychologists

like Wason thought that a rule that is a cause for serious logical troubles should be used for purposes of testing the rationality of human subjects.

An important caveat needs to be added here. What I have said so far applies only to examples like those mentioned. Gigerenzer and his colleagues, however, have developed a whole battery of rules of bounded rationality, so-called “fast and frugal heuristics” such as the “recognition heuristic”, the “fluency heuristic”, or “take the best”. This is the most important and novel part of the BR approach. These heuristics are adequate for problems characterized by inevitable uncertainty, very little information, or very little available solution time (e.g., Gigerenzer, Todd & the ABC Research Group, 1999). It is typical that these problems are often more realistic reasoning and decision problems than the textbook examples I have concentrated on so far. Logic or probability theory may not be very helpful in such cases. For instance, by using the recognition heuristic—“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”—we are able to draw highly accurate inferences about, say, the sizes of cities or winners in sports tournaments. Surprisingly, if we recognize only names of objects, we do better than if we have more information about them (Gigerenzer & Goldstein, 1996). Rules for such problems often depend on empirical knowledge, and are normatively valid only relative to certain domains or contexts. In the case of the recognition heuristic, this rule works if and insofar as recognition of city names correlates with the criterion. When there is no such correlation, judgments are wrong more often, and in such circumstances the heuristic should not be used.

I pointed out that the BR approach sometimes looks (and is characterized) as if it were about adequate reasoning *processes*, about determining which rules to *apply* to a given task or problem rather than about determining and justifying the *rules* themselves which we ought to follow in reasoning. Given the existence and normative validity of fast and frugal heuristics, it is clear that the program isn’t merely about such issues. It also develops for large numbers of problems and tasks, rules to be followed. But the fact that there are such norms of rationality does not show that there are no others.

This point brings me to my third and final consideration about limitations of the BR approach. It is perhaps tempting to infer from the foregoing arguments that not only is the psychology of reasoning deeply fragmented, but so is very concept of rationality. Our best instances of good reasoning may bear no more than a family resemblance to one another, some being guided by bounded norms, others by norms that are formal or strictly universal. However, this is probably incorrect, and not simply because we might be tempted to ask back: “Well, why is it that we call all of these different rules ‘norms of *rationality*’? They must have something in common after all!” Rather, I think that the two kinds of norms do, at least partly, play functionally different roles, and that they partly interlock. For instance, formal rules such as Bayes’ theorem or principles of optimization (such as maximizing

expected utility) continue to play an important role even within attempts to construct fast and frugal heuristics, namely as a normative standard. A heuristic of this kind can only be normatively recommendable if it *competes* successfully with such formal rules—if it leads to correct or convincing results at least as often as they do, or perhaps even outperforms them. Also, one cannot even formulate certain fast and frugal heuristics without using some basic concepts and tools of formal logic. If reasoners are to consciously use the recognition heuristic, they have to have a minimal grasp of the conditional, “if–then” structure of the heuristic. Other rules, such as “take the best” require the ability to master disjunction, and so on. In this way, some formal rules are built into heuristics. They are the cement that holds the building blocks of heuristics together. However, the BR approach also contains a fundamentally correct and important insight: because reasoning often has to proceed on the basis of very little information and large amounts of uncertainty, it makes little sense to expect logic or probability theory alone to be sufficient in a comprehensive normative theory of rationality. Heuristics can and should complement formal rules here.

Thus, the picture of rationality emerging from this might help overcoming the deep—and partly heated—debates in psychology discussed in this essay, at least in part. It does not do so by showing that both research programs have overplayed their cards, or that the differences are merely rhetorical while both programs share the same core assumptions about *how far* human beings are in fact rational (as argued by Samuels, Stich & Bishop 2002). My point concerns the different issue of the normative assumptions embodied in the programs. Without neglecting their basic differences, we can try to make them coherent by assigning different, interlocking functions to them. Certainly this will have to be discussed in closer detail. Could a more lasting peace emerge from this? I don’t know. But I hope the idea seems worth more research and collaboration between philosophy and psychology.¹³

Notes

1. For instance, they touch on debates about the possibilities and problems of naturalistic approaches to epistemology (e.g., Kitcher, 1992; Stein 1996, pp. 14–18). At the same time, epistemologists have so far taken little if any notice of the psychological debate about human rationality (exceptions being Bishop, 2008; Bishop & Trout, 2005; Goldman, 2008). This needs to change. Within the space of this essay, I cannot discuss the consequences of the “rationality wars” for naturalizing epistemology. But I hope to lay some groundwork for doing so.
2. The pessimism about human rationality that comes out of the HB approach has also worried naturalistic philosophers (e.g. Kornblith, 1992), though not all (e.g., not an earlier Stich, 1980; also, Bishop & Trout, 2005, try to exploit the HB approach positively for their own normative naturalism). Again, I must leave this interesting issue for another occasion.
3. There have been a number of critical discussions of the HB approach by philosophers as well, esp. Cohen (1981), Stein (1996), see also Botterill & Carruthers (1999), Cherniak (1986), Schumacher (2002). It would lead too far to discuss them here, however.

4. There are yet further objections from this camp against the HB approach. For instance, Gigerenzer (1996b) argues that the language of biases and heuristics is not really explanatory: talk of a “representative heuristic” or an “availability heuristic” are mere redescriptions of the phenomena. In other words, the HB approach does not really deliver *explanations*. Samuel, Stich and Bishop’s (2002) attempt to argue that the differences between the “heuristic and biases” and the “bounded rationality” approaches are merely superficial overlooks this important point.
5. As Gigerenzer remarks: “Recent studies using paraphrasing and protocols suggest that participants draw a variety of semantic inferences to make sense of the Linda problem: Some 10 to 20% seem to infer that and should be read as a conditional, and some 20 to 50% seem to infer that the alternative ‘Linda is a bank teller’ implies that she is not active in the feminist movement. . . . These semantic inferences can lead to choosing T&F rather than T” (1996a, p. 593). I have corrected the last sentence in the quotation. It originally reads: “These semantic inferences can lead to choosing T rather than T&F.” This is confused, as Gigerenzer has agreed in personal communication.
6. Gigerenzer often points out that there are serious debates within logic, statistics, and so on over which formal system is correct, which notion of probability adequate, etc. I cannot enter this side of his “reject the norm” argument here. See Vranas, 2000, 2001; Gigerenzer, 2001.
7. I ignore here whether it is best to call these inferences “semantic”.
8. Elimination of polysemy helps as well, but not as much as the use of frequentist language, as Gigerenzer has pointed out to me.
9. For instance, Cosmides and Tooby in their approach of an evolutionary psychology of rationality hold similar views. It should be mentioned that Gigerenzer’s BR approach is not committed to an evolutionary account of human rationality. For instance, it does not accept the strong modularity thesis that evolutionary psychologists maintain (Cosmides, 1989; Cosmides & Tooby, 1996). That the two views need to be distinguished is overlooked by Samuels, Stich and Bishop (2002).
10. Indeed, defenders of the HB approach have understood the concept of bounded rationality in this way (e.g., Gilovich, 1991; Kahneman, 2003).
11. E.g., “. . . on Kahneman and Tversky’s (1996) view of sound reasoning, the content of the Linda problem is irrelevant. All that counts are the terms *probable* and *and*, which the conjunction rule interprets in terms of mathematical probability and logical AND, respectively. In contrast, I believe that sound reasoning begins by investigating the content of a problem to infer what terms such as *probable* mean” (Gigerenzer, 1996a, p. 593).
12. I say “let us not generally confuse” them because there are cases of rules where the question of justification and application are not easily distinguishable. This holds especially for the “fast and frugal heuristics” of the BR approach. See Gigerenzer and Sturm (2011).
13. I thank Uljana Feest and Gerd Gigerenzer for valuable comments and criticisms. Cynthia Klohr made helpful suggestions for wording the text. Completion of this essay was supported by the Spanish Ministry for Science and Innovation, Reference number FFI 2008-01559/FISO.

References

- Bishop, M.A. & Trout, J.D. (2005) *Epistemology and the Psychology of Human Judgment* (New York: Oxford University Press).
- Bishop, M. (2008) “Reflections on a normative psychology”, in: A. Beckermann & Sven Walter (Eds.), *Philosophie: Grundlagen und Anwendungen*, pp. 249–62 (Paderborn: Mentis).
- Botterill, G. & Carruthers, P. (1999) “Reasoning and irrationality”, in: G. Botterill & P. Carruthers, *The Philosophy of Psychology*, pp. 105–30. (Cambridge: Cambridge University Press).

- Casscells, W., Schoenberger, A. & Graboys, T.B. (1978) "Interpretation by physicians of clinical laboratory results", *New England Journal of Medicine*, 299, pp. 999–1001.
- Cherniak, C. (1986) *Minimal Rationality* (Cambridge, MA: MIT Press).
- Cohen, L.J. (1981) "Can human irrationality be experimentally demonstrated?", *Behavioral and Brain Sciences*, 4, pp. 317–31 (comments and responses, pp. 331–59).
- Cosmides, L. (1989) "The logic of social exchange: Has natural selection shaped how humans reason?", *Cognition*, 31, pp. 187–276.
- Cosmides, L. & Tooby, J. (1996) "Are humans good intuitive statisticians after all?", *Cognition*, 58, pp. 1–73.
- Evans, J. St.B. & Over, D.E. (1996) *Rationality and Reasoning* (Hove: Psychology Press).
- Fiedler, K. (1988) "The dependence of the conjunction fallacy on subtle linguistic factors", *Psychological Research*, 50, pp. 123–29.
- Fodor, J. (2000) "Why we are so good at catching cheaters", *Cognition*, 75, pp. 29–32.
- Gigerenzer, G. (1991) "How to make cognitive illusions disappear: Beyond heuristics and biases", *European Review of Social Psychology*, 2, pp. 83–115.
- Gigerenzer, G. (1993): "The bounded rationality of probabilistic mental models", in: K.I. Manktelow & D.E. Over (Eds.), *Rationality: Psychological and Philosophical Perspectives*, pp. 284–313 (London: Routledge).
- Gigerenzer, G. (1996a) "On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky", *Psychological Review*, 103, pp. 592–96.
- Gigerenzer, G. (1996b) "Rationality: why social context matters", in: P. B. Baltes & Ursula M. Staudinger (Eds.), *Interactive Minds*, pp. 319–46 (New York: Cambridge University Press).
- Gigerenzer, G. (1998) "Psychological challenges for normative models", in: D. M. Gabbay & P. Smets (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, vol. 1., pp. 441–67 (Dordrecht: Kluwer).
- Gigerenzer, G. (2000) *Adaptive Thinking* (New York: Oxford University Press).
- Gigerenzer, G. (2001) "Content-blind norms, no norms, or good norms? A reply to Vranas", *Cognition*, 81, pp. 93–103.
- Gigerenzer, G. (2004) "The irrationality paradox", *Behavioral and Brain Sciences*, 27, pp. 336–38.
- Gigerenzer, G. & Goldstein, D. G. (1996) "Reasoning the fast and frugal way: Models of bounded rationality", *Psychological Review*, 103, pp. 650–69.
- Gigerenzer, G. & Hoffrage, U. (1995) "How to improve Bayesian reasoning without instruction: Frequency formats", *Psychological Review*, 102, pp. 684–704.
- Gigerenzer, G. & Hug, K. (1992) "Domain-specific reasoning: Social contracts, cheating, and perspective change", *Cognition*, 43, pp. 127–71.
- Gigerenzer, G. & Sturm, T. (2011) "How (far) can rationality be naturalized?" *Synthese*. DOI: 10.1007/s11229-011-0030-6.
- Gigerenzer, G., Todd, P.M. & the ABC Research Group (1999) *Simple Heuristics That Make Us Smart* (New York: Oxford University Press).
- Gilovich, T. (1991) *How We know What Isn't So: The Fallibility of Human Reason in Everyday Life* (New York: Free Press).
- Gilovich, T., Griffin, D., & Kahneman, D. (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment* (New York: Cambridge University Press).
- 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.), *Philosophie: Grundlagen und Anwendungen*, pp. 230–47 (Paderborn: Mentis).
- Grice, P. (1975) "Logic and conversation", in: P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts*, pp. 41–58 (New York: Academic).
- Grice, P. (2001) *Aspects of Reason* (Oxford: Clarendon Press).
- Hertwig, R. & Gigerenzer, G. (1999) "The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors", *Journal of Behavioral Decision Making*, 12, pp. 275–305.

- Kahneman, D., Slovic, P. & Tversky, A. (Eds.) (1982) *Judgment under Uncertainty: Heuristics and Biases* (New York: Cambridge University Press).
- Kahneman, D. & Tversky, A. (1996) “On the reality of cognitive illusions”, *Psychological Review*, 103, 582–91.
- Kahneman, D. (2003) “Maps of bounded rationality: Psychology for behavioral economics”, *American Economic Review*, 93, pp. 1449–75.
- Kant, I. [1781/1787] (1998): *Kritik der reinen Vernunft*, Jens Timmermann (Ed.) (Hamburg: Meiner).
- Kitcher, P. (1992) “The naturalists return”, *Philosophical Review*, 101, pp. 53–114.
- Lopes, L.L. (1991) “The rhetoric of irrationality”, *Theory & Psychology*, 1, pp. 65–82.
- Lopes, L.L. (1992) “Three misleading assumptions in the customary rhetoric of the bias literature”, *Theory & Psychology*, 2, pp. 231–36.
- Mellers, B. Hertwig, R. & Kahneman, D. (2001) “Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration”, *Psychological Science*, 12, pp. 269–75.
- Nisbett, R.B. & Borgida, E. (1975) “Attribution and the psychology of prediction”, *Journal of Personal and Social Psychology*, 32, pp. 932–43.
- Nisbett, R.B. & Ross, L. (1980) *Human Inference: Strategies and Shortcomings* (Englewood Cliffs, NJ: Prentice Hall).
- Oaksford, M. & Chater, N. (1994) “A rational analysis of the selection task as optimal data selection”, *Psychological Review*, 101, pp. 608–31.
- Piattelli-Palmarini, M. (1994) *Inevitable Illusions: How Mistakes of Reason Rule Our Minds* (New York: Wiley).
- Reisberg, D. (1997) *Cognition: Exploring the Science of the Mind* (New York: W. W. Norton).
- 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. 236–68 (New York: Oxford University Press).
- Schumacher, R. (2002) “Wie irrational können Personen sein?”, *Zeitschrift für philosophische Forschung*, 56, pp. 22–47.
- Sen, A. 1993: “Internal consistency of choice”, *Econometria*, 61, pp. 495–521.
- Simon, H. A. (1990) “Invariants of human behavior”, *Annual Review of Psychology*, 41, pp. 1–19.
- Stein, E. (1996) *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science* (Oxford: Clarendon).
- Stich, E. (1980) “Could man be an irrational animal? Some notes on the epistemology of rationality”, *Synthese*, 64, pp. 115–35.
- Sutherland, S. (1992) *Irrationality* (London: Pinker & Martin).
- Tversky, A. & Kahneman, D. (1971) “Belief in the law of small numbers,” *Psychological Bulletin*, 2, pp. 105–10. (Reprinted in Kahneman, Slovic & Tversky, 1982, pp. 23–31.)
- Tversky, A. & Kahneman, D. 1974: “Judgment under uncertainty: Heuristics and biases”, *Science*, 185, pp. 1124–31.
- Tversky, A. & Kahneman, D. (1983) “Extensional versus intuitive reasoning: Conjunction fallacy in probability judgment”, *Psychological Review*, 90, pp. 293–315.
- Tversky, A. & Kahneman, D. (1986) “Rational choice and the framing of decisions”, *The Journal of Business*, 59, pp. S251–S278.
- Vranas, P.B.M. (2000) “Gigerenzer’s normative critique of Kahneman and Tversky”, *Cognition*, 76, pp. 179–93.
- Vranas, P.B.M. (2001) “Single-case probabilities and content-neutral norms: A reply to Gigerenzer”, *Cognition*, 81, pp. 105–11.
- Wason, P.C. (1966) “Reasoning about a rule”, *Quarterly Journal of Experimental Psychology*, 20, pp. 273–81.
- Wason, P.C. (1983) “Realism and rationality in the selection task”, in: J. St. B. T. Evans (Ed.), *Thinking and Reasoning*, pp. 44–75 (London: Routledge & Kegan Paul).