

tuitive judgment, lest we lose it as a tool for criticizing and improving judgment. We cannot base normative theory on intuition, or on what intelligent people can be convinced of. Thus, in order to say that some judgment is biased, we need to do more than show that it is not reflectively endorsed, or even that it is “inconsistent.” The definition of consistency itself often presumes a normative theory (Baron 1994a).

Sunstein argues that “any ambitious theory is likely to be too contentious to serve as the benchmark for measuring moral truth” (sect. 3, para. 1). Why? Must we assume that consensus must be achieved in order to make progress? Must we accept the argument that “many people disagree” (or even “most scholars disagree”) as a killer argument against an ambitious normative theory such as utilitarianism, without examining the reasons for disagreement and whether they are responsive to the best arguments in favor of the theory? Some proponents of utilitarianism (an ambitious theory, for sure) think that some form of it can be derived logically from a useful analytic framework, such as the analysis of decisions into acts, states, and outcomes, with beliefs depending on states, and values depending on outcomes (e.g., Baron 2004; Broome 1991; Hare 1981; Kaplow & Shavell 2002). When opponents neglect our arguments, are we to simply give in? Give in to what?

Consider the problem that non-utilitarians have with distribution. Sunstein argues, “It is far from clear that a moderate utility loss to those at the bottom can be justified by a larger utility gain for many at the top” (sect. 3.1, para. 3). It is indeed far from clear intuitively, but the intuition that makes it unclear seems to be an overextension of a good utilitarian heuristic – that, other things being equal, the poor benefit more than the rich from a given good – to utility itself. Greene and Baron (2001) asked their subjects to evaluate distributions of utility, after making utility ratings of other goods (so that they had some idea of what utility was). The subjects showed declining marginal utility for the goods, as we would expect, but they showed just as much “declining marginal utility” for utility! This made them internally inconsistent. Greene and Baron argued that Rawls’s objection to the distributional consequences of utilitarianism is based on this overextension. Sunstein admits that such basic principles as Rawls’s difference principle could result from overextension of intuitive heuristics. My point here is that this kind of overextension may account for much of the difficulty of reaching consensus.

Some biases can be demonstrated by showing that they are not reflectively endorsed. They result from System I. Yet, many demonstrations of biases, such as omission bias, present the two cases to be compared (act and omission) adjacently, so that subjects have a chance to reflect. Many biases typically demonstrated with separated examples are also found with adjacent presentation (Frisch 1993). Some non-moral biases seem to resist even extensive argumentation, although they are clearly biases, such as Ellsberg’s ambiguity effect (Baron & Frisch 1994; see also Baron 2000b, pp. 268–73).

Possibly the most serious question that results from lack of consensus about normative theory is what prescriptive implications can be drawn from heuristics-and-biases research on policy judgments. It is difficult to impose utilitarianism on law and public policy when most people do not accept utilitarianism. Sunstein himself has faced this problem repeatedly and dealt with it creatively (Sunstein 2002; Sunstein & Thaler 2003).

Perhaps one other way to move forward without requiring consensus on utilitarianism (or any normative theory) is to focus on utilitarianism’s main feature, its focus on consequences. Sunstein comes close to this in his emphasis on “weak consequentialism.” Assume for a moment that the way to bring about the best consequences on the whole, to maximize utility, is to try to maximize utility. Then, any biases or heuristics that lead to different policies will make outcomes worse. The argument then becomes a conditional: If you are concerned about policies leading to consequences that are less good than they could be, then try to correct, or work around, the heuristics and biases that lead to suboptimal consequences (as argued in Baron 1998).

A possible problem with this argument is that the assumption it requires may be incorrect. It may be that we maximize utility only by trying to do something else, as Sunstein argues in the section on exotic cases. But the examples that make this argument plausible come mostly from personal behavior. In the domain of judgments about public policy, many other examples (such as those cited by Baron 1998) argue that the assumption is approximately correct: If we try harder to bring about good consequences, putting aside our nonconsequentialist intuitions, we might actually succeed.

Towards an intuitionist account of moral development

Karen Bartsch and Jennifer Cole Wright

Psychology Department, University of Wyoming, Laramie, WY 82071-3415.
bartsch@uwyo.edu narvik@uwyo.edu

Abstract: Sunstein’s characterization of moral blunders jointly indicts an intuitive process and the structure of heuristics. But intuitions need not lead to error, and the problems with moral heuristics apply also to moral principles. Accordingly, moral development may well involve more, rather than less, intuitive responsiveness. This suggests a novel trajectory for future research into the development of appropriate moral judgments.

Sunstein argues that, like other types of judgment, our moral judgments often employ heuristics (i.e., “mental short-cuts” or “rules-of-thumb”) that lead to blunders in the moral, legal, and political domains. Though we generally agree with his discussion, Sunstein’s intriguing portrayal of moral decision-making fails to adequately distinguish between two distinct aspects of the phenomenon. The first is the *process* by which moral heuristics are employed; the second is the *structure* of moral heuristics themselves.

Concerning *process*, Sunstein claims that moral heuristics are employed by the “intuitive system,” which is known for being “rapid, automatic, and effortless” (sect. 2.2, para. 3). We believe that Sunstein’s indictment of moral heuristics relies on a complaint against the intuitive system that may be unwarranted. First, intuitions are not necessarily grounded in heuristics. Second, we see no reason why “intuitive responsiveness” must lead to error. In fact, it may *protect* us from error; for, when adequately developed, intuitive responsiveness may reliably lead to appropriate moral judgments – a possibility we will expand on shortly.

Concerning *structure*, Sunstein’s discussion oscillates between at least two distinct types of moral heuristics: those of the traditional “rule-of-thumb” type (e.g., “do not tamper with nature”), and others more instinctive/affective in nature (e.g., the “outrage heuristic”). We will focus on the former. Sunstein argues that these heuristics are context-insensitive in a way that leads to unjustified, and sometimes dangerous, over-generalizations. We suggest that the structure of such heuristics is indistinguishable from the structure of moral principles. That is, we see no relevant structural difference between *heuristics* like “punish, and do not reward, betrayals of trust” and *principles* like “do not knowingly cause human death.” Consequently, *pace* Sunstein, the problem with heuristics is a problem with principles, as well.

Consider two of Sunstein’s moral heuristics: “people should not be permitted to engage in moral wrongdoing for a fee” and “do not tamper with nature.” As Sunstein points out, these heuristics are structurally blind to the many morally relevant details present in particular situations. But so are moral principles. Consider two well-known principles: “always keep your promise” (Kant 1948/1964) and “maximize utility” (Mill 1861/1957). Using these principles to guide one’s judgments can lead to moral blunders. For instance, keeping one’s promise is problematic in situations where it is morally appropriate to break the promise. Thus, the problems with applying a maxim to a complicated, contextualized problem

exist regardless of whether the maxim is a principle or a heuristic. Moreover, this is so whether such heuristics/principles are utilized rapidly, automatically, and effortlessly, or in deliberative reasoning. By themselves, they do not specify how/when they should be employed, how/when they admit of exceptions, how/when they ought to be used in conjunction with other heuristics/principles, and so on. In other words, there is a *gap* between how far reliance on moral heuristics/principles takes us and where we need to be in order to achieve appropriate moral judgments.

Some moral philosophers have answered this dilemma by positing a moral “sensitivity” (e.g., McDowell 1998; Railton 2000; Wiggins 1987/2002). But, what could this “sensitivity” be? We think that it may be the product of a well-functioning intuitive system, one that allows for rapid, automatic, effortless responsiveness without heuristics/principles. The question for moral psychology, then, becomes: how does one develop a well-functioning intuitive system, and thus moral maturity. If Sunstein’s critique of “rule-of-thumb” moral heuristics is correct (which we think it is), then the fact that heuristics and principles have identical structures suggests that this development cannot occur through the internalization of heuristics/principles alone.

Hubert and Stuart Dreyfus’s model of expertise (Dreyfus & Dreyfus 1986; 1991) might provide insight into the development of moral maturity. Their model suggests that heuristics/principles play a circumscribed role in moral development. If we consider moral maturity akin to other forms of expertise, then its development might be best characterized as a movement *away from*, rather than *towards*, moral judgments guided by heuristics/principles. In the expertise model, heuristics/principles are introduced early in development as basic rules that identify features recognizable without the benefit of experience (e.g., when learning chess, each piece is assigned a value and one is taught the rule “always exchange if the total value of pieces captured exceeds the value of pieces lost”). Reliance on such heuristics/principles is gradually replaced with procedural knowledge (i.e., know-how) gained through experience. Such knowledge leads to intuitive responsiveness. Intuitive responsiveness is the hallmark of expertise generally, because it enables rapid, automatic, effortless judgments in response to particular environmental contingencies. Importantly, such responsiveness is also reliably *appropriate*: this is what makes an expert an expert.

Examples might help. Just as the professional ski racer knows precisely how to adjust her posture to bring herself quickly around a steep turn; just as the concert pianist’s fingers move skillfully across the keys; just as the master chess player can play 5–10 second/move games without significant degradation in her performance – so, too, might the morally mature person simply *see* the moral relevance of particular situations and evaluate accordingly. Of course, we do not contend that intuitive responses are always correct, anymore than Sunstein maintains that they are always wrong. We simply wish to point out the need to treat the intuitive aspect of decision-making as a matter orthogonal to the issue of how heuristics/principles are applied, and to recognize that an intuitive response may in fact be characteristic of moral maturity.

In order to test the adequacy of the expertise model, researchers must gain insight into the development of moral maturity. This suggests a shift in emphasis from a focus on moral reasoning to the following sorts of questions: What kinds of activities lead to the development of moral know-how? What kinds of instruction/modeling are children morally responsive to? What kind of feedback best engenders moral sensitivity? Sunstein states that a primary goal of his article is to stimulate future research. We hope such research will include an exploration of these developmental issues, examining the potentially independent roles of intuitive processing and the application of heuristics/principles.

Neurobiology supports virtue theory on the role of heuristics in moral cognition

William D. Casebeer

National Security Affairs, Naval Postgraduate School, Monterey, CA 93943.

wdcasebe@nps.edu

<http://www.usafa.af.mil/dfpfa/CVs/Casebeer.html>

Abstract: Sunstein is right that poorly informed heuristics can influence moral judgment. His case could be strengthened by tightening neurobiologically plausible working definitions regarding what a heuristic is, considering a background moral theory that has more strength in wide reflective equilibrium than “weak consequentialism,” and systematically examining what naturalized virtue theory has to say about the role of heuristics in moral reasoning.

I agree with much of what Sunstein says about the role heuristics play in moral judgment and applaud his effort to make moral theorizing responsive to what is known about how human beings reason in concrete circumstances. The case for the existence of moral heuristics can be strengthened, however, by: (1) tightening working definitions regarding what constitutes a heuristic, with requisite sensitivity to their neurobiological underpinnings, (2) pressing for wide reflective equilibrium as we formulate our background “most plausible moral theory” so as to avoid charges of circularity, and (3) systematically considering what one major moral theory, a naturalized virtue theory, would say about the role heuristics play in moral cognition.

As we triangulate on a good theory about what a heuristic is, we should keep in mind the neurobiological substrates that constitute them. Despite the advanced state of play in the study of heuristics, it is difficult to articulate a framework that tells us with rigor just what they are. As Gigerenzer notes, one-word explanations can become surrogates for what should be richer psychological/neurobiological theories, saying of the representativeness, availability, and anchoring heuristics that “thirty years and many experiments later these three ‘heuristics’ remain vague and undefined, unspecified both with respect to the antecedent conditions that elicit (or suppress) them and also to the cognitive processes that underlie them” (Gigerenzer 2000, p. 290). In like vein, I worry whether candidate moral heuristics offered by Sunstein (e.g., “people should not be allowed to engage in moral wrongdoing for a fee,” “condemn as morally wrong things that outrage you,” or “do not tamper with natural processes for human reproduction”) really constitute *heuristics*. They sound much like candidates for potential moral *principles* (though not very promising ones). If I have independent reason (let’s stipulate this for the moment) for believing people ought not to be treated as a mere means, in what sense am I bringing a “heuristic” to bear on moral problem-solving when I apply Kant’s categorical imperative? Are heuristics present only in System I? Kantians would insist that they (not weak consequentialists) are in fact responding to the demands of the thoughtful, more detached, System II with their principles, as the deliverances of the categorical imperative are not automatized, influenced by emotion, subject to framing effects, and the like (or so they might maintain – this is probably false when we examine the neurobiological evidence). Allowing neurobiology to upwardly constrain theorizing about what is a heuristic will be useful, as minds and brains are always token-identical and perhaps even type-identical in some circumstances. If upon empirical investigation a heuristic has no plausible neurobiological substrate, nor any law-like connection to activation of evolved brain systems and architecture, that makes it a bad candidate for election to office. For reviews of the neurobiology of moral cognition, see Greene and Haidt (2002), Casebeer (2003a), and Casebeer and Churchland (2003).

My second concern is closely related. As a general methodological principle, theorizing in all domains should strive for consistency: at the very least, domains should be consistent, and in exemplary cases one domain might even be reduced to another (much scientific progress had been made in this way). “Wide,” rather than “narrow,” reflective equilibrium should be the norm.