

(Maio & Haddock, in press). When motivation and ability are higher, people tend to scrutinize the relevant arguments more carefully and disregard any heuristics that are unreliable or irrelevant. In this literature, “correctness,” though relevant, is downplayed. There is no algorithm for deciding whether someone spoke the truth – and the relevant standard by which something is deemed to be a heuristic is typically *other reasoning by the same user*, in situations of high engagement. This means it might even be possible for a particular statement to function as a heuristic on one occasion and as a valid premise on another (Kruglanski et al. 2004). In contrast, and more akin to the terminology of the naïve physics literature, Baron (e.g., 1993a; 1994a) introduces the term “moral heuristic” for the rules that constitute our “naïve morality” (e.g., Baron 1993a). Examples include “it is wrong to hurt some people for the benefit of others” or “harmful commissions are worse than harmful omissions.” Though similar in appearance to persuasion heuristics, the status of these rules as heuristics is not determined by processing context. Another perspective on moral reasoning emphasises that moral judgment might be achieved through two separate cognitive systems: an intuitive system and a reasoning system (Haidt 2001). The “intuitive system,” which Sunstein equates with heuristics, is characterised as fast and effortless; its processing is unintentional, typically inaccessible to awareness, and involves parallel processing and pattern matching. For moral judgment, this intuitive system additionally involves emotion (Greene & Haidt 2002; Haidt 2001).

Crucially, these related but distinct notions of the term “heuristic” all require different kinds of evidence. Evidence for “cognitive” heuristics in the Tversky and Kahneman sense requires patterns of judgment that deviate in the predicted fashion from some standard of correctness (see also, Kahneman & Tversky 1996). By contrast, the Gigerenzer sense requires evidence of the opposite, namely, accuracy. Whether processing was deliberate or automatic, conscious or unconscious, or involves affect, is, at least in the first instance, unimportant for both (though see now, Kahneman 2002; Kahneman & Frederick 2002) and evidence for heuristics in problem-solving was even derived largely from verbal protocols of reasoners describing their thinking out loud. By contrast, evidence for “persuasion” heuristics requires demonstration that their use is influenced by motivation and ability. Finally, evidence for the “intuitive system” is virtually orthogonal to that required for “cognitive” heuristics: standards of correctness are irrelevant, and processing characteristics are all important.

Sunstein’s article seems to simultaneously endorse all of the above uses, in that “deviations from correctness” and “output from System 1” and “adaptiveness” are variously emphasised. However, most of the examples given are content rules, part of our naïve morals in Baron’s sense. That is, they are “moral principles that are generally sound, and even quite useful, but that work poorly in some cases” (sect. 5.1.1, para. 2). Evidence for these principles as heuristics is then supplied by describing a case for which they seemingly lead to an “incorrect” answer.

In order to evaluate Sunstein’s proposal we turn to consideration of legal systems as complex systems explicating our sense of right and wrong. Setting aside the vexed issue of absolute standards of correctness, one finds that it is a property of *all* legal rules and principles that eventually cases will emerge for which their application suggests an undesired outcome. Real legal systems try to minimize this problem through a proliferation of rules of different scope, whereby the system is supplemented with further rules defining exceptions and exceptions to exceptions. Unanticipated exceptions will nevertheless arise. In other words, legal norms are inherently *defeasible* (see e.g., Bankowski et al. 1995). Seen in this light, there is little point in calling a moral content statement a “heuristic” simply because it can and eventually will give rise to an unwanted “overgeneralization.”

This suggests to us that the term “moral heuristic” would better be limited to *processes*. The target article provides no real evidence to this effect. However, we have, for example, recently found intriguing effects of typicality. In several experiments (Frost

et al., in preparation), we asked participants to analyse their reasons for the value of equality in a typical context (gender discrimination) and an atypical context (handedness discrimination). Results indicated that participants who generated reasons in a typical context later acted in a more egalitarian manner than participants who generated reasons in the atypical context, despite listing similar numbers of reasons for the value and being equally confident in their reasons. Here, issues of correctness seem unproblematic, as participants see no reason why their behaviour on the same task should differ according to exposure to previous material. At the same time, it seems safe to assume that the workings of this particular typicality effect are entirely opaque to participants. In short, the concept of a moral heuristic might yet prove useful in explaining moral judgment and behaviour, but only if it is about more than particular content rules or principles, which are prone to exception.

Invisible fences of the moral domain

Jonathan Haidt

Department of Psychology, University of Virginia, Charlottesville, VA 22904.
haidt@virginia.edu <http://www.people.virginia.edu/~jdh6n/>

Abstract: Crossing the border into the moral domain changes moral thinking in two ways: (1) the facts at hand become “anthropocentric” facts not easily open to revision, and (2) moral reasoning is often the servant of moral intuitions, making it difficult for people to challenge their own intuitions. Sunstein’s argument is sound, but policy makers are likely to resist.

Look at it from Bin Laden’s point of view. For years the United States had been. . . . don’t worry, I’m not going to finish the sentence. I can’t. I study morality and I know that terrorism is driven largely by moral commitments. Yet, every time I try to understand Bin Laden, or Hitler, or political leaders with whom I strongly disagree, I feel a kind of invisible fence (the kind used for suburban dogs) giving me a warning shock, saying “don’t go there, don’t even think about empathizing.” In contrast, I can roam freely around the Linda problem, the Asian Disease problem, and the visual illusions that I use to show my Psych 101 students how perceptual heuristics can sometimes misfire. It can be difficult to look at a probability problem or a perceptual illusion in a different way, but it is never dangerous or painful.

Sunstein’s effort to bring the well-developed tools of research on heuristics into moral psychology is welcome and well done. His emphasis on “System I” processes in the moral homunculus is consistent with recent emphases on the role of emotion and intuition in moral judgment (Damasio 1994; Greene et al. 2001; Haidt 2001). However, the moral domain is a weird and treacherous world in which objects change their weights and rivers flow uphill. Or at very least, minds that worked in one way on non-moral problems suddenly start working differently when moral concerns are introduced. Here I discuss two such differences which I believe can be integrated into Sunstein’s approach, giving us a fuller and more social picture of the workings of moral heuristics.

1. Moral truths are anthropocentric truths. Sunstein contrasts the moral domain with the “domain of facts,” suggesting that moral truths are not facts, but this is not quite right. A useful distinction can be made between two kinds of facts – anthropocentric and non-anthropocentric (Wiggins 1987). Non-anthropocentric facts are those that do not depend for their truth on the way the human mind is constituted. Facts about the physical world and mathematical truths are true regardless of what we happen to think about them, and they would presumably be true for any intelligent species that came to our solar system to inspect them. But our judgments about beauty, humor, and morality are factual judgments too. They are judgments about anthropocentric truths – truths that are true only because of the kinds of minds that we

happen to have, and the cultural worlds within which our minds developed. When we give an A+ to one paper and a D to another we are asserting that one paper really is better than the other, within our academic community, although we might not expect intelligent extraterrestrials to agree with us.

Anthropocentric truths arise within communities, and they then do much of the work of marking out the limits of those communities. But even within the realm of anthropocentric truths, moral facts are especially potent. Groups can usually tolerate a diversity of beliefs about beauty and comedy, but moral diversity is much more damaging (Haidt et al. 2003). One cannot even coherently want moral diversity. For example, if a person says, "I believe that women should have the right to choose, but I would prefer that there be a diversity of opinions on that matter," then that person treats abortion rights as a taste, not as a moral issue. Foundational beliefs, such as taking the bible as the literal word of God, or the idea that the world is full of victims of oppression who must not be blamed for their fate, become sacralized, and those who question them risk becoming pariahs. Many moral heuristics may have this sacred character for some groups (e.g., don't play God, don't knowingly kill anyone, don't have sex with your family members, don't blame victims), so questioning them, even in special cases where they don't really apply, is likely to meet with resistance and even outrage. The problem with moral heuristics is not that there is no fact of the matter with which to compare them; rather, it is that there are many (anthropocentric) facts of the matter, and it is hard to get people to question their anthropocentric moral facts.

2. In the moral domain, System II is often a slave. In Sunstein's analysis, System II (reasoning) either opposes System I (by reaching a conclusion that the homunculus opposes) or it sits back and does nothing while System I spits out its heuristic conclusion. But, in any domain in which strong motivations are at work, reasoning often becomes the "slave of the passions," as David Hume put it. We can sometimes see this process at work for non-anthropocentric facts, as when students struggle to find reasons to explain how Linda is more likely to be a bank teller active in the feminist movement than to be a bank teller. But many people are able to reason their way to a solution, and there is often a moment of insight in which System II triumphs and people understand their error. Not so for moral disputes. I have now interviewed several hundred people about taboo violations such as consensual safe sex between an adult brother and sister, and I have never yet seen a person say "Oh, I see! I had this strong gut feeling that it was wrong, but now that I understand that no child can result from the union, I realize that I was mistaken." More typically, people struggle valiantly to find some reason why even in this special case the brother and sister should not have sex. We can therefore expect a lot less help from System II in challenging moral heuristics than we get from it in challenging non-moral heuristics. In fact, whenever moral emotions are engaged, as they often are when anthropocentric facts are challenged, we can expect to find the System I homunculus ordering System II to man the ramparts and fight off persuasion.

I think these two differences make a difference. I applaud Sunstein's call for distrust of moral heuristics when considering unusual or difficult cases. And I expect that many people will agree with him, as I do, in the abstract. But when it comes time to make policy decisions about abortion, euthanasia, cloning, or any other difficult issue, don't be surprised when politicians and policy makers refuse to cross their invisible fences, or when they attack those who ask them to do so.

ACKNOWLEDGMENTS

I thank the John Templeton Foundation for its generous support of my work.

Sunstein's heuristics provide insufficient descriptive and explanatory adequacy

Marc D. Hauser

Departments of Psychology, Organismic and Evolutionary Biology, and Biological Anthropology, Harvard University, Cambridge, MA 02138.

Mdhauser@Wjh.Harvard.edu <http://www.Wjh.Harvard.edu/~Mnkylab>

Abstract: In considering a domain of knowledge – language, music, mathematics, or morality – it is necessary to derive principles that can describe the mature state and explain how an individual reaches this state. Although Sunstein's heuristics go some way toward a description of our moral sense, it is not clear that they are at the right level of description, and as stated, they provide no guidelines for looking at the acquisition process – the problem of explanatory adequacy.

Consider the human language faculty. When we generate sentences, or comprehend them, we do so effortlessly. Our capacity to both understand what others say and to generate new prose is boundless. The way to make sense of this capacity is by appealing to a dedicated faculty of the mind, a system that contains a repository of computational resources for building an externalized language. For each individual, the language they construct, both over their lifetime, as well as on a moment to moment basis, represents the output of a complicated series of interfaces between the computational resources dedicated to language, on the one hand, and interactions with other mind internal–external factors, on the other hand. Linguists interested in the underlying principles that can account for what a mature speaker of a language knows are studying the descriptive principles of the system.

One of the early mysteries surrounding this approach to language was the observation that young children are able to both generate surprisingly sophisticated sentences and comprehend them in the absence of relevant input. This observation led in part to the hypothesis that our species is innately equipped with a universal grammar, a set of principles and parameters that not only enables the capacity to build a natural language, but also constrains the range of possible languages. The now rich description of the principles and parameters in play early on in development provides a sense of the explanatory adequacy of this field.

In this commentary, I make use of the importance of descriptive and explanatory adequacy in characterizing a domain of knowledge, as well as the tie in to language, to evaluate Sunstein's discussion of our moral psychology. I first describe the shortcomings of the moral heuristics position and then provide a sketch of an alternative which builds on an analogy with language (Dwyer 1999; 2004; Harman 1999; Hauser, in press; Hauser et al., in press; Jackendoff 2004; Mikhail 2000; Mikhail et al. 2002; Rawls 1971; Smith 1759/1976).

Sunstein wants to show that heuristics play a significant role in moral, legal, and political spheres, and that sometimes they generate inappropriate judgments. As stated, it is hard to imagine that anyone would disagree with these claims. Those who thought hard about common sense morality, beginning with Hutcheson and Shaftesbury, recognized that we often apply general rules of thumb in cases of moral conflict and, as Hume importantly recognized, funnel these rules through an emotional filter that guides our actions. What have always been the primary challenges to these views include our ability to understand where our common sense intuitions come from, what their representational content is, the extent to which they are consciously available principles as opposed to unconscious and inaccessible, how children alight upon them in the course of attaining a mature moral faculty, and the degree to which they facilitate or detract from our interests in normative or prescriptive principles aimed at a just world. Concerning the latter, the interest has always been a concern with how our intuitions or heuristics about right and wrong interface with more formal and explicit policies, whether they are the unstated social norms of a hunter-gatherer society or the legal doctrine of our founding fathers. So, on a general level, there is not much new in