Response to Nichols and Katz
Richard Joyce

To reject a false theory on the basis of an unsound argument is, in my opinion, as much an intellectual sin as to embrace a false theory. Thus, although I am no fan of any particular form of moral rationalism—and, indeed, on occasion have gone out of my way to criticize it—when rationalism is assailed for faulty reasons I find myself in the curious position of leaping to its defense (which goes to show that in philosophy it isn't the case that one's enemy's enemy is one's friend). This puts me at something of a dialectical disadvantage, since my "defense" of moral rationalism has strict limits: I will defend it against *specific kinds* of criticism, but I have no interest in defending it *simpliciter*. It is important to bear in mind that what is in dispute between Shaun Nichols and myself is not the truth or falsity of moral rationalism, but rather *what kind of evidence bears on the matter.*

Concerning Psychological Rationalism Nichols and I have no argument. Concerning Conceptual Rationalism there are two sources of disagreement. First, Nichols acknowledges that the experimental design that he originally employed was inadequate to refute Michael Smith's brand of Conceptual Rationalism, since the subjects were not asked whether the imaginary psychopath was "irrational." But Nichols doubts that this would have made much difference to the outcome, pointing to the widespread view of Satan, who is popularly conceived of as "the rational amoralist *par excellence*." The problem with this response is that "rational" is, in the vernacular, a fairly amorphous and indeterminate term, concerning which (I'll wager, in advance of experimental evidence) ordinary people have vague, conflicting, pliable, and quite possibly downright confused opinions. There is certainly *a sense* in which Satan is considered rational. He's good at math, could get an A+ in *Logic 101* without breaking a sweat, and is extremely effective at means-end reasoning. But surely we should acknowledge another sense in which ordinary people may hesitate about Satan's rationality: They may suspect that there is something irrational about Satan's *ends* (his desire to inflict endless suffering, etc.); they may be tempted by the thought that the way that Satan has failed to see things "correctly" (i.e., the way God sees them) counts as some kind of rational failure; indeed, they may simply think that Satan's evil character is evidence of his irrationality. Whether there is any philosophical sense to be made of the idea that practical ends can be assessed as rational/irrational is something that can, of course, be doubted; it is a longstanding problem in moral philosophy. But that there exists a body of widespread firm (but possibly nebulous) opinions favoring this idea is something about which, it seems to me, there can be little doubt. After all, all those neo-Kantians who devote their energies to making sense of the idea must be picking up on *something* in the popular consciousness. I predict that ordinary subjects will, when questioned about rationality (e.g., when asked to assess Satan in this manner), be easily manipulated either to assert or deny claims of the form "X is rational" depending on the framing of the query. (On framing effects in general, see Kahneman and Tversky 1979).

This brings us nicely to the second, more general, disagreement between Nichols and myself over Conceptual Rationalism: namely, the extent to which consulting the opinion of ordinary speakers (using questionnaires and the like) is likely to cast interesting light on conceptual content. Michael Smith is skeptical on the grounds that conceptual content is best conceived of as a systematization of the platitudes surrounding a term—where these platitudes are

"descriptions of the inferential and judgemental dispositions of those who have mastery of the term." Note Smith's reference to "*dispositions*"—which he understands as a kind of knowledge *how* to use a term. In Smith's opinion, what we end up with when we systematize platitudes— that is, when we analyze a concept—is knowledge-that about knowledge-how (Smith 1994: 38). And the key characteristic of knowledge-how, of course, is that those who possess it—even those who are masters of it—may be quite unable to articulate it. (How many of us can describe with any precision what we do with our bodies when we descend a flight of stairs?) When platitudes concern knowledge-*how*, then, *pace* Nichols, one should not "rather expect the folk to recognize their own platitudes," any more than one should expect a concert pianist to be able to articulate how her eyes scan the musical notation. Of course, there's no harm in asking the folk for their opinion, and there's no reason to be "dismissive" of their reply (just as it may be interesting to see what the concert pianist has to say about her eye movements), but what really needs to be directly examined is how the folk use their terms in everyday interaction and decision-making (just as what we really need to do in order to find out about a pianist's eye movements is simply to observe them in action).

My principal goal in the lead paper was to emphasize how the traditional philosophical form of moral rationalism concerns practical *justification*. And such justification, I claimed, is unlikely to be tractable by *a posteriori* methods. This latter claim was something I put forward as more-or-less obvious, without submitting much by way of argument. Nichols responds by offering some interesting thoughts concerning how justification may in fact be affected by empirical discoveries. His argument is based on the general claim that empirical discoveries about the genealogy of certain beliefs can show those beliefs to be epistemically unjustified.[1] More specifically: He argues that Justificatory Rationalism depends for its plausibility on the acceptance of a certain intuitive premise which, when its psychological sources become known, will be epistemically undermined. Yet more specifically: He argues that Justificatory Rationalism—Peter Singer's version, to be precise—depends on the acceptance of a "Justification Principle," but that our acceptance of this intuitive principle in fact derives from "nonrational affective mechanisms" in such a way as to cast our acceptance of this principle into doubt.

The general point that beliefs can be epistemically undermined on the basis of their origins is one with which I have much sympathy; indeed, I have advocated the idea myself on several occasions (Joyce 2001, 2006*a*, 2006*b*), including in the closing paragraphs of the target paper. I also agree that having their source in "nonrational affective mechanisms" may be one way that beliefs (or "intuitions") may be debunked (see also Sinnott-Armstrong 2005). However, I doubt very much that Justificatory Rationalists see their argument as turning on a premise that enjoys the backing of nothing more than its "seeming intuitive." Indeed, I would go so far as to say that the urge to avoid building their theory on so shaky a basis is one of the Justificatory Rationalists' guiding intellectual motivations. The challenge for the Justificatory Rationalist, then, is to provide an argument—including a premise if not along the lines of Nichols' Justification Principle then at least playing the same role—that appeals to more than just "powerful intuition"—that does, in fact, command the assent of any rational observer. At this point in the debate, it must be stressed, I really must shift to speaking vicariously—outlining what I assume the rationalist's response will be rather than pressing an argument that I have any genuine faith in. And it must also be confessed that I am a poor proxy for a real committed moral rationalist, for in fact I have a longstanding puzzlement about this point in Justificatory Rationalist's

---

[1]. Or *render* those (formerly justified) beliefs epistemically unjustified (depending on one's epistemological tastes).

argument (in just the way Nichols has), and I too harbor grave doubts that the rationalist's argument will ultimately work. (See my comments on Katz below.) But I at least recognize that the Justificatory Rationalist is *endeavoring* to base the argument on something firmer than "intuition."

Many Justificatory Rationalist arguments, for example, ultimately boil down to some sort of "Treat like as like" principle: If two things are qualitatively the same in a certain respect, then it is incorrect to treat them as if they were not alike in that respect. This principle is supposed to extend to the guidance of our treatment not just of persons, but of cats, books, and rocks. The Justificatory Rationalist thinks that denying such a principle is not merely counter-intuitive, but palpably incoherent. Consider, by comparison, the logical rule of *modus ponens*. To what do such logical rules owe their central place in deductive reasoning? This very good question is not one concerning which I would here hazard an answer; it is enough for our purposes to make the negative observation that the status of *modus ponens* does not derive simply from its "intuitive obviousness"—or, at the very least, if the rule *is* based on intuition it is in a manner that does not undermine our commitment to it. The Justificatory Rationalist hopes to invest "Treat like as like" with an analogous status: not derived simply from "intuitive obviousness"—or, at the very least, if based in intuition then in a manner that does not undermine our commitment to it. The fact that "Treat like as like" is supposed to be a cool-headed basic truth of practical rationality—holding as much for our treatment of rocks as for persons—casts doubt on the plausibility of the conjecture that our acceptance of it is grounded, ultimately and problematically, in the activity of emotional mechanisms. The aspiration to prescribe fairness on the basis of rationally compelling premises, rather than relying on any sentimental appeal that fairness may happen to have, is precisely what distinguishes Justificatory Rationalism from its rivals.

Leonard Katz offers a very different set of critical comments, focused less on the arguments presented in my contribution to this volume, and more on a broader philosophical program that I have advocated in other books and articles. That program, in essence, is to argue that the most justified position to take with respect to moral discourse and thought is analogous to the atheist's view of religious discourse and thought. This "error theoretic" view holds that although our moral judgments aim at the truth, they systematically fail to secure it. Space does not permit me to attempt to mount a defense this form of moral skepticism; I must rest content with making a few brief comments and at least locating the crux of the disagreement between Katz and myself.

It should be noted that although one could, in principle, be a moral error theorist *by implication*—either because one endorses a radical global error theory (thus being skeptical of morality along with modality, colors, other minds, cats and dogs, etc.), or because one endorses an error theory about all normative phenomena—typically the moral error theorist thinks that there is something *especially* problematic about morality, and does not harbor the same doubts about normativity in general. The moral error theorist usually allows that we can still deliberate about how to act, she thinks that we can still make sense of actions harming or advancing our own welfare (and others' welfare), and thus she thinks that we can continue to make sense of various kinds of non-moral "ought"s, such as prudential ones.[2] Thus the moral error theorist can without embarrassment assert a claim like "One ought not harm others," so long as it is clear that it is not a *moral* "ought" that is being employed. (In the same way, an atheist can assert that one ought not covet thy neighbor's wife, so long as it clear that this isn't an "…according to God" prescription.) Holding a moral error theoretic position does not imply any degree of tolerance for

---

[2]. For detailed discussion of the moral error theorist's attitude towards prudential normativity, see Joyce (forthcoming). Certain passages hereabouts are taken straight from this paper.

those actions we generally abhor on moral grounds. Although the moral error theorist will deny (when pressed in all seriousness) that the Nazis' actions were morally wrong, she also denies that they were morally right or morally permissible; she denies that they were morally *anything*. This does not prevent her from despising and opposing the Nazis' actions as vehemently as anyone else.

It must also be remembered that there are many possible routes to a moral error theory, and one mustn't assume that the metaethical position is refuted if one argumentative strategy in its favor falters. Perhaps the error theorist thinks that for something to be morally bad (e.g.) would imply or presuppose that human actions enjoy a kind of unrestricted *autonomy*, while thinking that in fact the universe supplies no such autonomy (see Haji 1998, 2003). Perhaps she thinks that for something to be morally bad would imply or presuppose a kind of inescapable, authoritative imperative against pursuing that thing, while thinking that in fact the universe supplies no such imperatives (Mackie 1977; Joyce 2001). Perhaps she thinks that for something to be morally bad would imply or presuppose that human moral attitudes manifest a kind of uniformity, while thinking that in fact attitudes do not converge (Burgess 1978). Perhaps she thinks that there exists no phenomenon whose explanation requires that the property of moral badness be instantiated, while thinking that explanatory redundancy is good ground for disbelief (Hinckfuss 1987). Perhaps she thinks that tracing the history of the concept *moral badness* back to its origins reveals a basis in supernatural and magical forces and bonds—a defective metaphysical framework outside which the concept makes no sense (Hägerström 1953). Perhaps she thinks all these things and more besides. Perhaps she is impressed by a number of little or medium-sized considerations against morality—none of which by itself would ground an error theory, but all of which together constitute A Big Problem. (In what follows, however, for ease of reference I will speak as if the error theorist accuses morality of just one monolithic failure.)

Most straightforward objections to the moral error theoretic position fall into one of two categories. The opponent may acknowledge that the putatively problematic attribute that the error theorist assigns to morality really is problematic, but deny that this attribute is an essential component of morality; a normative framework stripped of the troublesome element will still count as a morality. Alternatively, the opponent may accept that the putatively problematic attribute is a non-negotiable component of anything deserving the name "morality," but deny that it really is problematic. So, for example, if the error theorist is claiming that moral properties require a kind of pure autonomy which the universe does not supply, then one type of opponent will insist that morality requires nothing of the sort, while another will insist that the universe does indeed contain such autonomy.

Katz's response paper includes elements of both kinds of objection. First, he admits that "some parts of natural human morality may rest on illusion," but thinks that there is plenty that remains unscathed after the error theorist's attack. In particular, he thinks that the types of objection that I have raised in favor of moral skepticism do not undermine the existence of *objective hedonic values*, and much of morality can be built upon these values. Second, Katz focuses on a specific "problematic attribute" of morality on which I have concentrated earlier efforts (especially in my 2001 book)—namely, that morality centrally depends on our having desire-transcendent practical reasons—and he argues that this attribute is in fact not problematic at all; objective hedonic values can ground desire-independent reasons.

Katz's second strategy is clearly his dominant one (the first is, in any case, hard to assess, since he leaves it unclear which bits of morality he is willing to concede are "illusory"), so it will be the focus of what little remains of this rejoinder. The general pattern of this dialectic is one that any error theorist will face over and over again: A "contender" property (or cluster of properties) is brought forward, for which it is claimed (A) that it is instantiated by the world and

(B) that it satisfies the criteria sufficient to count as a "morality." The error theorist's commitment involves her denying either (A) or (B); usually (B). By comparison, think again of the atheist. Imagine him being assailed by opponents offering accounts of God designed to dispel his doubts: "Well, God is just *love*, and you believe in love, don't you?—so you believe in God." The atheist's answer, obviously, is that it is simply wrong to literally identify God with love, and were someone really to insist on doing so, then she would be using either the term "God" or the term "love" (or, I suppose, the word "is") in an eccentric and stipulative manner. To believe in the thing denoted by the vernacular English word "God" involves a range of unusual ontological commitments that believing in love does not.

Though Katz's opposition to moral skepticism is nowhere near as crass as the "God is love" example, the broad structure of my defense is the same: To believe in the things denoted by the English words "moral rightness," "evil," "moral badness," "moral duty," etc. involves a range of commitments that believing in pleasure and pain will not supply. The way that I have formulated the debate in the past (Joyce 2001) has invited Katz to think that what is crucial to our dispute is whether pleasure and pain will supply *reasons*. I am now not so sure that this is a necessary, or the most perspicacious, way of proceeding, but in any case I won't change tack now. The questions that need to be put to this particular "contender" are the following: Does pain (say) ground values? Does pain ground reasons? Does pain ground objective reasons? Does pain ground moral normativity? It seems pretty clear that Katz conceives of these questions roughly as a building sequence, such that if any one of them receives a "No" this will upset the plausibility of the subsequent questions receiving a positive answer. My view is that the sequence might be reasonably interrupted at *any* point, though space doesn't allow the development of supporting arguments. I will instead close by cutting right to the heart of the problem.

Let me grant for the sake of argument that my pain is a (negative) value for me. Let me also grant that my pain supplies me with reasons for acting. And I'll also grant (again, purely *arguendo*) that these reasons are (in some sense of the term) objective. The trouble begins by wondering why my pain is a value for, and why it supplies reasons to, *anyone other than me*. There is a gaping chasm between "My pain gives me reason to act" and "My pain gives you reason to act." But without this chasm being bridged, the values and reasons that Katz manages to squeeze into the world will simply not ground a *morality*, for moral normativity centrally involves *other persons* being loci of value and imposers of duties. Though I do not wish to exclude the possibility of self-promoting moral values and duties, I would argue adamantly that a normative system consisting *wholly* of such values and duties—in which one's duties to others is entirely contingent on what effect the discharge of those duties has on one's own pleasure states, and in which the value of others depends completely on their bearing on one's own hedonic welfare—does not count as a *moral* system.

Katz would, I think, concur. He recognizes that if pleasure and pain are to undergird the whole moral system, it is vital that one individual's pleasure and pain episodes can ground *someone else's* values and reasons—someone, that is, who isn't the least bit sympathetic to the individual in question. But on this absolutely vital step he is content to offer no more than "cf. Nagel 1986." The use of "cf." leaves us guessing the extent to which Katz endorses Thomas Nagel's views, but we are forced to assume at least some sympathy, for there is nothing else offered by Katz that promises to accomplish this crucial step in the argument. But though Nagel's views are subtle and influential, in my opinion his attempt to show that an episode of pain—irrespective of *whose* pain—provides *anybody* with reason to alleviate that pain, must count as a notable and noble failure. (See Dancy 1988; Mack 1989; Carlson 1990.) To demonstrate this I would really need to go through the relevant passages of Nagel's works claim

by claim, which I cannot do here. I must be satisfied to observe that on this point of contention hinges the whole of Katz's case against my brand of moral skepticism, and on this point we disagree.

REFERENCES:

Burgess, J.P. 1978. "Against ethics." Unpublished manuscript. Department of Philosophy, Princeton University.

Carlson, G. 1990. "Pain and the quantum leap to agent-neutral value." *Ethics* 100: 363-367.

Dancy, J. 1988. "Contemplating one's Nagel." *Philosophical Books* 29: 1-15.

Hägerström, A. 1953. *Inquiries into the Nature of Law and Morals*. Stockholm: Almqvist & Wiksell.

Haji, I. 1998. *Moral Appraisability*. Oxford: Oxford University Press.

Haji, I. 2003. *Deontic Morality and Control.* Cambridge: Cambridge University Press.

Hinckfuss, I. 1987. "The moral society: Its structure and effects." *Discussion Papers in Environmental Philosophy* 16. Canberra: Philosophy Program (RSSS), Australian National University.

Joyce, R. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.

Joyce, R. 2006*a*. *The Evolution of Morality*. Cambridge, MA.: MIT Press.

Joyce, R. 2006*b*. "Metaethics and the empirical sciences." In J. Kennett and P. Gerrans (eds.), *Philosophical Explorations* 9: 133-148.

Joyce, R. Forthcoming. "Morality, schmorality." In P. Bloomfield (ed.) *Morality and Self-Interest*. Oxford: Oxford University Press.

Kahneman, D. and Tversky, A. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica* 47.

Mack, E. 1989. "Against agent-neutral value." *Reason Papers* 14: 76-85

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin Books.

Nagel, T. 1986. *The View from Nowhere*. Oxford: Oxford University Press.

Sinnott-Armstrong, W. 2005. "Moral intuitionism meets empirical psychology." In T. Horgan & M. Timmons (eds.), *Metaethics After Moore*. New York: Oxford University Press.

Smith, M. (1994). *The Moral Problem,* Oxford: Blackwell.