



Evolution and Impartiality

Author(s): Guy Kahane

Source: *Ethics*, Vol. 124, No. 2 (January 2014), pp. 327-341

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/673433>

Accessed: 17/02/2015 03:25

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

<http://www.jstor.org>

Evolution and Impartiality*

Guy Kahane

Katarzyna de Lazari-Radek and Peter Singer argue that evolutionary considerations can resolve Sidgwick's dualism of practical reason because such considerations debunk moral views that give weight to self-interested or partial considerations but cannot threaten the principle of universal benevolence. I argue that even if we grant these claims, this appeal to evolution is ultimately self-defeating. De Lazari-Radek and Singer face a dilemma. Either their evolutionary argument against partial morality succeeds, but then we need to also give up our conviction that suffering is bad; or there is a way to defend this conviction, but then their argument against partiality fails. Utilitarians, I suggest, should resist the temptation to appeal to evolutionary debunking arguments.

I. EVOLUTIONARY DEBUNKING EXPLANATIONS

In many of its forms, utilitarianism is in direct tension with many commonsense moral beliefs and intuitions, such as the common belief that partiality to self, dear and near, is often justified. Utilitarians try to explain, and to explain away, the persistence of such anti-utilitarian intuitions. One way to undermine such intuitions is to offer debunking explanations of their source: explanations that show them to have their source in biasing historical influences, such as now discredited religious or metaphysical views.¹ In some of his recent work, Peter Singer has shifted the focus further back in time, to the evolutionary origins of common intuitions that conflict with utilitarianism.² The idea is that utilitarians can cast

* I am very grateful to the Editor and two referees for helpful suggestions. Work on this article was supported by grant from the Wellcome Trust (WT087208MF).

1. See, e.g., Peter Singer, "Sidgwick and Reflective Equilibrium," *Monist* 58 (1974): 490–517.

2. Peter Singer, "Ethics and Intuitions," *Journal of Ethics* 9 (2005): 331–52.

Ethics 124 (January 2014): 327–341

© 2014 by The University of Chicago. All rights reserved. 0014-1704/2014/12402-0005\$10.00

doubt on intuitive opposition to utilitarianism by tracing its source to evolutionary pressures that bear no relation to the moral truth.

As several authors have pointed out, this is a risky strategy, since, even if it succeeds in undermining nonutilitarian views, it might also end up undermining Singer's own utilitarian view.³ After all, it seems that even though utilitarianism involves a radical departure from common-sense moral views, it nevertheless builds on basic moral dispositions and sentiments that are also likely to have been selected by evolution. This result is also implied by the more sweeping global evolutionary debunking arguments that have been developed by Richard Joyce and Sharon Street, arguments that target all of our moral beliefs.⁴

In their recent paper, Katarzyna de Lazari-Radek and Peter Singer make intriguing observations about relevant remarks in Sidgwick that anticipate the recent debate about evolutionary debunking, and about how this debate bears on Sidgwick's dualism of practical reason.⁵ But there is a good sense in which their main aim is to respond to the challenge I have just outlined: they attempt to show both that universal benevolence (UB), the core utilitarian claim that "The good of one individual is of no more importance, from the point of view of the universe, than the good of any other"⁶ is in fact immune to evolutionary debunking,⁷ and that its various partialist competitors, including rational ego-

3. Folke Tersman, "The Reliability of Moral Intuitions: A Challenge from Neuroscience," *Australasian Journal of Philosophy* 86 (2008): 389–405; Guy Kahane, "Evolutionary Debunking Arguments," *Noûs* 45 (2011): 103–25. Similar points are made in Neil Levy, "Cognitive Scientific Challenges to Morality," *Philosophical Psychology* 19 (2006): 567–87; Selim Berker, "The Normative Insignificance of Neuroscience," *Philosophy and Public Affairs* 37 (2009): 293–329.

4. Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT Press, 2007), chap. 6; Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127 (2006): 109–66.

5. Katarzyna de Lazari-Radek and Peter Singer, "The Objectivity of Ethics and the Unity of Practical Reason," *Ethics* 123 (2012): 9–31. Although Sidgwick's remarks are suggestive, I think that de Lazari-Radek and Singer overstate Sidgwick's prescience (*ibid.*, 12). First, Sidgwick is of course right to say that we cannot debunk beliefs just by showing that they have some causal explanation. But debunking challenges only arise when that explanation shows some beliefs to have their source in processes that aren't truth-tracking; de Lazari-Radek and Singer's entire argument presupposes this point. When discussing Sidgwick's remarks, they also appear to overlook the point that to show in this way that beliefs actually have an unreliable source is not the same as demanding a noncircular positive account of reliability. Finally, following Sidgwick, de Lazari-Radek and Singer suggest that because ethics and science operate in distinct domains, we cannot infer the falsity of all moral claims from anything about the causal origins of our beliefs. But debunking explanations don't aim to show that some (or all) beliefs are false, only that they lack justification. I'll return to this point.

6. *Ibid.*, 27.

7. De Lazari-Radek and Singer concede that UB may be on a par, in this respect, with other unqualifiedly impartial moral views. To simplify the discussion, I will follow them in setting this aside in what follows.

ism, aren't. This is how evolutionary considerations are supposed to dissolve Sidgwick's dualism of practical reason.

De Lazari-Radek and Singer's main argument is based on two claims about what our evolutionary history can, and can't, explain: they argue that it can explain our self-centered and other partial tendencies, but that it cannot explain belief in UB.

These are controversial empirical claims, and we should not overestimate our ability to confidently assess their plausibility. And it is hard to fail to notice that opinion as to whether some moral view is vulnerable to evolutionary explanation often tends to suspiciously align with an author's prior normative and metaethical commitments.⁸ This is why, although I have some serious reservations about the way that de Lazari-Radek and Singer defend these empirical claims,⁹ the main criticism I will be developing here will not directly challenge them or rely on any similarly controversial empirical speculation.

Instead, I will argue that even if we grant these claims, the authors fail to address the worry that utilitarian appeals to evolutionary debunking are ultimately self-undermining. As we shall see, their position faces an unwelcome dilemma. If their evolutionary debunking argument against partiality succeeds, then a parallel argument should also force us to give up our belief that suffering is bad and other core evaluative beliefs about well-being. But if they want to resist this unpleasant conclusion, then they must give up their argument against impartiality.

Before we proceed, let me first briefly clarify what such debunking explanations can achieve. At various points, de Lazari-Radek and Singer

8. We should be skeptical not only of speculative evolutionary just so stories, but also of comforting just *not* so stories—confident assertions that this or that simply can't be explained in evolutionary terms.

9. For example, de Lazari-Radek and Singer's debunking argument against rational egoism seems to focus on the wrong target. They correctly note that it is plausible that there is an evolutionary basis to the belief that we have reason to promote our self-interest. This normative claim, however, is not only compatible with UB, but actually implied by it. A more appropriate target is the common belief that we have reason to give priority to our own good, and that of our kin. And it is plausible that this view has an evolutionary basis (see Street, "Darwinian Dilemma," 115). But rational egoism is the claim that we have reason to promote *only* our own good. This claim goes beyond the above claim about self-interest, and conflicts with commonsense partial morality. It is far less obvious that this peculiar view can be given an evolutionary explanation. Evolutionary theorists have long argued that straight egoism (even if one attempts to conceal it) would be a poor reproductive strategy in the tightly knit small communities in which our distant ancestors lived (see, e.g., Robert Frank, *Passions within Reason* [New York: Norton, 1988]). If anything, it's likely that there is an evolutionary explanation for the common rejection of rational egoism. For similar criticism, see Roger Crisp's remarks in the vigorous online exchange at *Pea Soup* (<http://peasoup.typepad.com/peasoup/2012/12/ethics-discussions-at-pea-soup-katarzyna-de-lazari-radek-and-peter-singer-the-objectivity-of-ethics-1.html>), which contains further interesting criticism of these two claims.

write as if evolutionary explanations can debunk various principles.¹⁰ But this is misleading. Evolutionary explanations tell us nothing about the truth or falsity of moral claims or principles. What they can do is undermine the justification of beliefs in some moral claims or principles, by showing that they have their source in processes that are not truth-tracking.¹¹ When de Lazari-Radek and Singer say that evolutionary considerations debunk rational egoism, they must mean that it debunks the intuitive justification that Sidgwick and others hold belief in rational egoism to possess.

II. IMPARTIALITY, WELL-BEING, AND THE BADNESS OF SUFFERING

De Lazari-Radek and Singer develop their argument in some detail. But toward the end of their paper, they acknowledge a large gap in the argument. The principle of UB tells us to treat the good of all individuals in an impartial way but, as they admit, it “tells us nothing about what this good may be.” And without some theory of well-being to plug into it, UB is “empty of content.”¹²

The problem is that, as the authors also admit, there is a serious worry that our beliefs about well-being are also easy targets for evolutionary debunking. As I noted in an earlier paper, many of our evaluative beliefs about well-being, including the beliefs that pleasure is good and pain is bad, are some of the most obvious candidates for evolutionary debunking.¹³

This is a worry that de Lazari-Radek and Singer do not really address. But it means that even if UB was really immune to evolutionary debunking—even if we had positive grounds for thinking that belief in UB has its source in reliable general rational capacities—this would achieve little. For if our beliefs about well-being still get debunked, then UB can do no normative work, and offers no guidance to action.

I wrote above that we should temper our confidence in many evolutionary explanations of our normative beliefs. But if *any* normative be-

10. See, e.g., de Lazari-Radek and Singer, “Objectivity of Ethics,” 23.

11. Kahane, “Evolutionary Debunking Arguments.”

12. De Lazari-Radek and Singer, “Objectivity of Ethics,” 28.

13. Kahane, “Evolutionary Debunking Arguments,” 18. Here and in what follows, I will take well-being to be an evaluative notion, referring to what Sidgwick calls “an individual’s good.” There may also be a purely descriptive sense of well-being, which leaves it an open question whether what affects someone’s well-being (in that sense) also affects their good, or generates any reasons for action. Evolutionary explanations need not challenge beliefs about well-being in this descriptive sense (if there are any)—but such beliefs are irrelevant here.

lief can be given such an explanation, it is the universal (or near-universal) conviction that pain is bad.¹⁴ As Street writes,

it is of course no mystery whatsoever, from an evolutionary point of view, why we and the other animals came to take the sensations associated with bodily conditions such as [cuts and burns] to count in favour of what would avoid, lessen, or stop them rather than in favor of what would bring about and intensify them. One need only imagine the reproductive prospects of a creature who relished and sought after the sensations of its bones breaking and its tissues tearing; just think how many descendants such a creature would leave in comparison to those who happened to abhor and avoid such sensations.¹⁵

Unlike the evolutionary argument for UB, these claims require no elaborate speculations about our evolutionary history, or about the reliability of our capacity for reason. If evolution has had *any* role in shaping our evaluative beliefs, it would be hard to deny, to put it mildly, that evolution has played a key role in disposing us to see pain as bad, as something we have strong reason to avoid and minimize.

But this means that our evaluative beliefs about pain and pleasure, and about many other aspects of well-being, clearly fail to meet one of de Lazari-Radek and Singer's conditions for thinking that an intuition is reliable, namely "the absence of a plausible explanation of the intuition as the outcome of an evolutionary or other non-truth-tracking process."¹⁶ The authors make no attempt to resist the above claims about the evolutionary source of our beliefs about well-being. In fact, they concede that these are plausible claims.

14. To my knowledge, Sidgwick never considers this worry about our evaluative beliefs about hedonic states—not even when he cites Herbert Spencer's evolutionary account of the origins of our concern with pleasure and pain (Henry Sidgwick, *The Methods of Ethics*, 7th ed. [London: Macmillan, 1907], bk. II, VI, sec. 3). This omission suggests that Sidgwick wasn't really anticipating the current discussion of evolutionary debunking arguments to the extent suggested by de Lazari-Radek and Singer.

15. See Street, "Darwinian Dilemma," 150. Notice that Street's lengthy discussion of pain (144ff.) involves various complications that are redundant for my purposes here. Street's ultimate aim is to argue that pain's badness is best interpreted along antirealist lines, as having its source in our negative evaluative attitude to the sensation of pain. We are now asking only whether our belief in pain's badness would be vulnerable to evolutionary debunking, if we assume, as de Lazari-Radek and Singer do, an objective conception of practical reason. Such a worry could be raised even if we were confident that pain's badness cannot be explained along antirealist lines. We can thus avoid entering into difficult questions about the nature of pain, and how it relates to our unreflective evaluative attitudes.

16. Ibid., 26. As de Lazari-Radek and Singer state this condition, it is merely a ground for not thinking that some intuition is reliable. From this it doesn't immediately follow that it is ground for thinking the intuition to be *unreliable*. But in their argument against par-

This is a remarkable admission. It would mean that even if their argument is entirely successful, it merely supports the conditional conclusion that if something is good for individuals, we should treat it with complete impartiality. Needless to say, this conclusion would be of no interest if nothing is good for individuals. Now, as noted earlier, debunking arguments cannot directly show that nothing is good for individuals. But they can undermine our justification for believing that any specific thing is good for individuals. And if we have no clue what matters for well-being, then UB is rendered entirely idle. If de Lazari-Radek and Singer's argument is successful, then in a formal sense, it would still allow us to resist Street's sweeping skeptical conclusion: there would be at least one moral belief we will be justified in holding, even if we assume an objective conception of practical reason.¹⁷ But in every other respect, the upshot would be just the same, since UB on its own generates no reasons for action. Order has hardly been restored to the Cosmos of Duty. There may no longer be Chaos, but that is only because the Cosmos of Duty is utterly empty.¹⁸

De Lazari-Radek and Singer do offer several brief remarks that function as damage control. After noting this challenge, they respond that they will limit themselves to pointing out that "if no theory of well-being or intrinsic value were immune to a debunking explanation, this would show only that no theory could be preferred over others on the ground that it alone cannot be debunked. It could not show that no theory of well-being is true."¹⁹ But this response is inadequate. We can immediately set aside that last remark about truth. To repeat, debunking arguments tell us nothing about the truth of some view. They only undermine its justification, a point that de Lazari-Radek and Singer occasionally forget. But this point about truth is irrelevant here. The skeptical outcome I described above would result if the justification of all of our beliefs about well-being was undermined in this way; that some theory about well-being might still be true won't help at all.

tiality, and in their remarks about the act/omission distinction, de Lazari-Radek and Singer clearly assume this stronger claim. This more plausible stronger claim is also assumed by much of the discussion of evolutionary debunking arguments; see, e.g., Joyce, *Evolution of Morality*; Street, "Darwinian Dilemma"; Tersman, "Reliability of Moral Intuitions"; and Kahane, "Evolutionary Debunking Arguments."

17. Street of course doesn't endorse this skeptical conclusion, but takes it as ground for rejecting evaluative realism. I'll return to this below.

18. Strictly speaking, debunking arguments leave open the possibility that the Cosmos of Duty is full; they just cast doubt on our ability to see what is in it. This point, however, also holds of de Lazari-Radek and Singer's own argument: even if successful, their debunking explanation of rational egoism cannot directly show that it is false, only that we have no reason to believe in it.

19. De Lazari-Radek and Singer, "Objectivity of Ethics," 21.

Consider next de Lazari-Radek and Singer's suggestion that, if no theory of well-being is immune to debunking explanation, then this would merely serve to leave us where we started, with whatever other reasons we previously had to prefer one theory to another. This puzzling remark again seems to misunderstand what debunking arguments achieve, if successful. It's not as if *not* being debunked adds some further positive justification to some theory, on top of all the other justifications for and against various theories. Rather, being debunked defeats the current justification for a theory, unless some alternative source of justification can be found.²⁰ If we disagree about why our cult leader is so great, but then discover that we believe that he is great (and great in different ways) only because we were brainwashed, then what we should conclude is that we have no reason to think that the cult leader is great in any way, not that, because all of our beliefs about this are equally unreliable, we can resume our disagreement.

If Street's general debunking argument is successful then, coupled with the assumption of evaluative realism, it will leave us with general evaluative skepticism. It won't merely level the epistemic playing field. In the same way, if all our evaluative beliefs about well-being were debunked, this would leave us with what we can call "well-being skepticism"—with no reason to believe that pain, pleasure, desire-satisfaction, love, deep personal relationships, knowledge, moral virtue, accomplishment, spiritual experience, and so forth, matter to anyone's good. And this, as we have seen, would have the same upshot as general evaluative skepticism, even if belief in UB is immune to debunking.

That would be the upshot if no theory of well-being or intrinsic value was immune to debunking explanation. But there is another thing that is odd about the above remarks by de Lazari-Radek and Singer. They move far too quickly from the worry that some of our beliefs about well-being might be given debunking explanations, to discussing the possibility that all of these beliefs might be debunked in this way.

De Lazari-Radek and Singer seem to forget their own complaint against Street: that, in arguing that all of our moral beliefs can be debunked, Street ignores the fact that some moral views (such as those justifying greater concern for our kin) seem far more susceptible to evolutionary explanation than others (such as those requiring complete impartiality).²¹ But surely the same is also true in the case of well-being. Some views about well-being seem far more susceptible to debunking explanation than others.

20. De Lazari-Radek and Singer write here as if a debunking explanation amounts to a further first-order consideration against some view. But such explanations aren't first-order reasons counting against some view, but undermining defeaters of some or all of the first-order reasons that we had previously assumed to support the view.

21. *Ibid.*, 19.

Our evaluative beliefs about pain and pleasure are perhaps the easiest to explain in evolutionary terms. It will be hard, at best, to find a serious evolutionary theorist who would deny this. These hedonic beliefs fall at one end of the scale, with beliefs about the value of fulfilling our desires probably coming a fairly close second. But it is also easy to find beliefs that fall at the other end of the scale. These would include the views that the good life consists of ascetic contemplation of deep philosophical truths, or celibate spiritual communion with God, or a kind of Nietzschean perfectionist aestheticism (which might even revel in pain), and so forth.

If we accept, as de Lazari-Radek and Singer do, the assumption that if an intuition has its source in a non-truth-tracking process, then this debunks its justificatory force, then many of our core evaluative beliefs about well-being are likely to be debunked. But other, more revisionary accounts of well-being might survive—or at least survive to fight another day.

If this somewhat less bleak result holds up,²² then UB would still have some concrete normative content and thus offer genuine guidance to action. This result will also be compatible with a rather peculiar form of utilitarianism. But I doubt that de Lazari-Radek and Singer will see this as a great improvement over general evaluative skepticism about well-being. If our core evaluative beliefs about well-being, including our beliefs about suffering, were debunked in this way, then one immediate casualty is utilitarianism in nearly all of its classical and contemporary forms, including, of course, those favored by Sidgwick and Singer. And it goes without saying that the assumption that suffering is bad, and pleasure good, is absolutely central to many of the controversial normative views that Singer has famously defended over many decades.

Utilitarianism is often viewed as an extremely counterintuitive view; many find Singer's normative views troubling, even repugnant. But if we take the goal of purging all evolutionary influence from our normative views seriously enough, we will end up with a view that is so radically divorced from common sense, and so distant from any familiar ethical theory, that, by comparison, Singer's own utilitarianism will seem almost like old-fashioned common sense. (How odd it is to assume that, if we purge our ethical beliefs of all evolutionary influence, we will end up with a normative view that precisely coincides with one of the ethical theories that some have defended prior to such purging.)

I very much doubt that de Lazari-Radek and Singer would be willing to seriously entertain an ethical theory which, although utilitarian in form, would deny suffering any intrinsic badness, nor see anything in-

22. Though of course further inquiry may reveal that even these more counterintuitive views ultimately have their source in evolutionary pressures (or other "off-track" causal influences—they may, for example, serve some useful social function), in the same way that further inquiry may reveal that UB has such a source.

trinsically good in happiness. I would join them in rejecting such a view, as would many others. But notice that their only plausible ground for rejecting such a revisionary view would have to be their intuitive conviction that suffering must be bad, and that if some argument leads to a contrary conclusion, it must be mistaken. If this is de Lazari-Radek and Singer's response then, ironically, they would be making exactly the kind of conservative appeal to intuition, in the face of strong countervailing considerations, that Singer has criticized throughout his career.

III. CAN'T HAVE IT BOTH WAYS: WHY DE LAZARI-RADEK AND SINGER CAN HAVE EITHER UNIVERSAL BENEVOLENCE OR PAIN'S BADNESS, BUT NOT BOTH

It might be objected that even if the question of well-being raises a serious problem, this is a problem that lies beyond the scope of de Lazari-Radek and Singer's main argument. They have defended UB from evolutionary challenges, while debunking partiality. If their argument is successful, this is already a major achievement. The question of well-being is a further issue, to be dealt with separately.

This objection fails. As we shall now see, the question of well-being isn't a further issue to be dealt with later. It is intertwined in multiple ways with de Lazari-Radek and Singer's defense of UB. To see this, consider the following possible responses to the evolutionary challenge to our beliefs about suffering's badness.

Rejecting the Evolutionary Explanation?

One obvious way to resist the challenge is to deny that our belief in pain's badness is tainted with any kind of evolutionary influence. As I noted above, this move seems rather hopeless. But for our purposes here, a weaker claim is sufficient: that it is highly implausible that our belief in pain's badness doesn't have an evolutionary source, yet our belief in various forms of partiality does.

In any event, a closer inspection of de Lazari-Radek and Singer's argument against partiality reveals that it leaves them with little space to deny that an evolutionary explanation can be given of our core beliefs about well-being. This is because, like UB, rational egoism and other partial normative views are also empty of content, and give no normative guidance, without some account of well-being. Thus, belief that one should promote one's good leads to concrete behavior *only* conjoined with some beliefs about what it *is* that will promote one's good. But this means that if these partial normative views are based on evaluative dispositions selected by evolution—as required by de Lazari-Radek and Singer's argument—then what evolution selected must have been a package containing both these self-centered and partial dispositions and some view

of well-being. And that view of well-being no doubt must have included the relevant core beliefs about pain, pleasure, death, and the like. So their argument against partiality really implicitly assumes that these core beliefs about well-being can be given an adaptive evolutionary explanation.²³

Appealing to Our Reliable General Rational Capacities?

De Lazari-Radek and Singer argue that evolution has provided us with general rational capacities that allow us to track objective normative truths. If this argument is successful, it suggests another way to defend our normative beliefs about suffering, and other aspects of well-being. By using this capacity for reason, we may find that although evolution has pushed us to believe that pain is bad, it turns out (by a lucky coincidence) that this belief is also independently rationally plausible.²⁴ This move is left open by their argument. The problem is that the same response can also be given by the defenders of partiality, since many would similarly claim that their belief in partiality is also independently rationally plausible. And if to that de Lazari-Radek and Singer reply that it merely *seems* to us that our intuitions about partiality or self-interest have their source in pure reason—that it is really just *evolution* we are hearing—then again, the same can also be said about our conviction that pain is bad.²⁵

Appealing to Our Substantive Normative Convictions?

In a footnote, de Lazari-Radek and Singer refer to Knut Skarsaune's response to Street, which aims to show that we can answer her challenge to evaluative realism in the case of pain's badness.²⁶ Although de Lazari-Radek and Singer don't explicitly endorse this response, they appear to think that it offers one way of addressing the evolutionary challenge to our beliefs about well-being.

23. This is not meant as a general claim about the way evolution works. Evolutionary pressures can of course also favor general capacities (such as the rational capacities discussed in the next section), if these lead to advantageous behavior. All I am claiming is that, without familiar core beliefs about well-being, there would be little or no evolutionary advantage in the more abstract belief that one should give priority to one's own well-being, or that of one's kin, over the well-being of others.

24. Derek Parfit appears to endorse this move in his *On What Matters* (Oxford: Oxford University Press, 2011), vol. 2.

25. De Lazari-Radek and Singer may reply that whereas belief in pain's badness does survive rational reflection, beliefs in justified partiality, or in rational egoism, do not. But Sidgwick famously thought that belief in rational egoism is as rationally plausible as belief in UB, and, as a matter of fact, many at least claim that their belief in partiality survives rational reflection. These claims may be mistaken. But if they can be shown to be mistaken on purely a priori grounds, de Lazari-Radek and Singer would not need to appeal to evolution in the first place.

26. Knut Skarsaune, "Darwin and Moral Realism: Survival of the Iffiest," *Philosophical Studies* 152 (2011): 229–43.

However, Skarsaune doesn't offer a direct defense of our belief in pain's badness. Rather, he points out that if pain really is bad, then we can offer a "third factor" account of why, although evolutionary pressures aren't truth-tracking, they could still dispose us to form true beliefs about pain's badness (notice that Skarsaune simply assumes that these beliefs about pain are due to such evolutionary pressures). Skarsaune's next move is to argue that if pain isn't really bad, then while this would indeed mean that our evaluative beliefs are systematically mistaken, this implication would hold independently of any evolutionary argument and that, moreover, evaluative realism is perfectly compatible with such a skeptical possibility.

Now, whether or not Skarsaune's "iffy" defense of realism against Street's challenge succeeds, it clearly can't help de Lazari-Radek and Singer, since it doesn't offer a defense of our conviction that pain is bad. To offer such a defense, we would need to add the further premise that it is legitimate for us to appeal to our deep evaluative convictions when we try to resist the evolutionary challenge to realism—in this case, that in trying to resist this challenge, it is legitimate for us to assume that pain is bad.

This is a controversial move. It can seem question-begging, yet, as other defenders of a third-factor approach point out, if it is illegitimate, then it is hard to see how we can avoid an implausible general radical skepticism, since it would be similarly impossible to show that our perceptual and mathematical beliefs are reliable without assuming some core convictions from these domains.²⁷

For our purposes, however, there is no need to resolve this issue. It is enough to note that if this move is legitimate, then it should also apply in other contexts. In particular, it should also apply with respect to strong convictions about justified partiality, including, of course, those relating to our own self-interest. So whether or not this line of response offers a way out, it's a way out that also lets the partialist escape, and in principle could get around any putative debunking challenge to a normative conviction that we find antecedently plausible. So the challenge remains: either evolutionary debunking arguments work as advertised, but their scope is so broad that they lead to general evaluative skepticism or something close, or they can be resisted fairly easily, in which

27. See David Enoch, "The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It," *Philosophical Studies* 148 (2010): 413–38; Erik Wielenberg, "On the Evolutionary Debunking of Morality," *Ethics* 120 (2010): 441–64. See also Roger White, "You Just Believe That Because. . ." *Philosophical Perspectives* 24 (2010): 573–615. For some doubts about this move, see Richard Joyce, "Evolution, Truth-Tracking, and Moral Skepticism," in *Problems of Goodness: New Essays on Metaethics*, ed. B. Reichardt (Bonn: Bernstein Verlag, forthcoming); Russ Shafer-Landau, "Evolutionary Debunking, Moral Realism and Moral Knowledge," *Journal of Ethics and Social Philosophy* 7 (2012): 1–37.

case they cannot support UB or any other significant normative view (nor, of course, any grand metaethical conclusions).

There is a broader issue in the background. De Lazari-Radek and Singer present Sidgwick as having basically anticipated, and adequately responded to, the skeptical challenge presented by evolutionary explanation. They thus allow themselves to ignore much of the growing recent literature attempting to respond to the evolutionary challenge put forward by Street and others. But de Lazari-Radek and Singer accept one key assumption of Street's: that if a belief (or intuition) has its causal source in a process that isn't truth-tracking, then this belief can be debunked. As we just saw, this assumption is rejected by proponents of so-called third-factor responses to evolutionary challenges.²⁸ But if we can reject this assumption, then de Lazari-Radek and Singer's argument against partiality immediately loses its force.

Rejecting Realism?

De Lazari-Radek and Singer present themselves as defending the objectivity of practical reason, but they actually say little about metaethical issues, or about how Street's evolutionary challenge is supposed to bear on them. Now Street assumes that the evolutionary challenge works only if we assumed the truth of evaluative realism (or, in de Lazari-Radek and Singer's terms, the objectivity of practical reason).²⁹ This is an assumption that they clearly accept.

Street's evolutionary argument for antirealism is, briefly, as follows:

- A. The conjunction of evaluative realism and our evolutionary history implies general moral skepticism.
- B. Moral skepticism is false.

Therefore—on the assumption that we aren't willing to give up the theory of evolution³⁰—it follows that

- C. We must reject evaluative realism.

Now if de Lazari-Radek and Singer's argument is successful, then we (can) know at least one moral claim, UB, even if realism/objectivism is true, and premise A is therefore false. The above argument against realism can be blocked.

28. See Enoch, "The Epistemological Challenge"; Wielenberg, "On The Evolutionary Debunking of Morality."

29. Though, as Street clarifies in later work, her understanding of 'realism' is rather broad, and even encompasses some forms of noncognitivism/quasi-realism.

30. For someone willing to question evolutionary theory on such grounds, see Thomas Nagel, *Mind and Cosmos* (Oxford: Oxford University Press, 2012).

But it is easy enough to broaden this argumentative strategy. In Street's version, it relies on the assumption that moral skepticism is false—that we know at least some evaluative truths. But it is just as likely that this claim either relies on, or can be extended to, the claim that we at least know certain specific evaluative claims (or even that we know certain evaluative claims with certainty), and therefore have reason to reject views that imply that we don't actually know these claims. And if there are evaluative claims that we know with certainty, one of these is surely that suffering is bad. The plausibility of a metaethical view that implied either that pain isn't bad, or that we don't know that pain is bad, would suffer greatly.

Now *if* we accept de Lazari-Radek and Singer's evolutionary argument, then we can both reject partiality and hold onto evaluative realism (or the objectivity of practical reason). My argument so far was that if we do so, then we must also give up the belief that suffering is bad. I've considered several ways of resisting this last result. The problem was that they all seem to also undermine de Lazari-Radek and Singer's argument against partiality. But an alternative way to resist this result is simply to give up evaluative realism. Once we give up realism, then the debunking force of evolutionary explanation is defused,³¹ and the evolutionary debunking argument against our belief in pain's badness loses its force. The problem, again, is that then the parallel argument against partiality also loses its force.

Worse, although utilitarianism, UB, and other forms of unqualified impartiality are logically compatible with antirealist views, they are extremely hard to defend once we accept a broadly Humean antirealist metaethics of the kind favored by Street, on which moral truth is constructed out of our subjective attitudes and sentiments, including, presumably, the sentiments and intuitions underlying much opposition to utilitarianism.³²

So which of these packages is more plausible: realism and UB without pain's badness, or antirealism and pain's badness, without UB? I think the latter would be far more attractive to many. In fact, I wonder whether it won't be significantly more attractive even to de Lazari-Radek and Singer.³³

31. Street, "Darwinian Dilemma"; Kahane, "Evolutionary Debunking Arguments." Joyce argues that some forms of antirealism may still be in principle vulnerable to evolutionary debunking (see his "Evolution, Truth-Tracking, and Moral Skepticism"). This however doesn't bear on the present argument, which only requires that at least some forms of antirealism can block a particular debunking argument.

32. This point is a major background theme of Bernard Williams's work. See, e.g., his *Moral Luck* (Cambridge: Cambridge University Press, 1981).

33. In personal communication, Singer once suggested to me that if the evolutionary debunking argument against realism can't be resisted, he would be inclined to respond by

So not only is the problem of well-being a truly damaging gap in their argument, but, contrary to the impression given by de Lazari-Radek and Singer, it is not even an independent issue—a further question to be dealt with after they have successfully defended UB. For how we respond to the problem of well-being also directly bears on (and is almost certain to undermine) the force and plausibility of their argument for impartiality.³⁴

In conclusion, de Lazari-Radek and Singer face a dilemma. If their argument against self-interest, partiality, and rational egoism works, then it also works against core beliefs about well-being. If it doesn't work against these core beliefs, then it also fails to undermine self-interest, partiality, and rational egoism.³⁵

endorsing some form of antirealism. The considerations above suggest that in this case he might also be forced to give up a strictly impartial view, or at the very least one key argument for such a view.

34. In passing, de Lazari-Radek and Singer remark that they will leave open the question of "what kind of method could be used to determine what is ultimately good" ("Objectivity of Ethics," 27–28). It's worth briefly considering how this question might bear on the challenge I have pressed. There seem to be three main ways by which we could try to answer questions about the good of individuals. One common focus of debate about well-being is intuitions about particular cases (see, e.g., Derek Parfit, *Reasons and Persons* [Oxford: Clarendon, 1984], app. I). Views about well-being might also be driven by more abstract intuitions, such as that which de Lazari-Radek and Singer take to underlie UB. On either approach, our intuitions about pain's badness would be open to evolutionary explanation. A third approach might see claims about what constitutes well-being as analytic truths, written into the very concept of well-being. Perhaps this would block the worry that our belief in pain's badness fails to track some independent reality. But this approach is implausible. Disagreements about what constitutes well-being seem as substantive as other disagreements in normative ethics. And although there is wide agreement about the claim that pain is bad, this claim isn't an analytic truth—its denial is perfectly intelligible. Nor could merely linguistic convention endow pain with its horrid badness, or otherwise remove it. In any event, it seems ad hoc and unstable for de Lazari-Radek and Singer to hold that questions about morality concern substantive, objective matters, yet questions about well-being merely concern analytic truths. Let me finally note that, with most contemporary authors, Sidgwick took questions about the ultimate good to be substantive questions, to be answered by reflection on our "intuitive judgments" and "a comprehensive comparison of the ordinary judgments of mankind" (*Methods of Ethics*, bk. III, XIV, sec. 5).

35. An objection: "You have been overlooking an important difference between our belief in the badness of pain (or goodness of pleasure) and other normative beliefs, such as belief about justified partiality. We have direct introspective access to pain's badness. We can *just see* that it's bad. No evolutionary story could challenge that." In reply, notice first that this move introduces a distinct route to reliable normative belief, one that is utterly different from (and perhaps in tension with) de Lazari-Radek and Singer's appeal to our general rational capacities. It also assumes that pain's badness is due to its experiential character. This is a view that I find plausible, but that many reject—including, on a common reading, Sidgwick himself (see *Methods of Ethics*, bk. II, II, sec. 2). We also need to distinguish having introspective access to our being in pain, or even to what it is like to be in pain (which, let us assume, is the ground of pain's badness)—both of which are entirely compatible with debunking explanations of our belief in pain's badness—and having intro-

Utilitarians often find stubborn intuitive resistance to their view frustrating. Evolutionary debunking arguments seem to offer an enticing new way to finally undermine this opposition. But this is a temptation that utilitarians would be wise to resist. They should join forces with the many authors who have recently tried to block such evolutionary debunking arguments, and pin their hope on defending utilitarianism on other grounds. Otherwise, they risk a kind of epistemic suicide.³⁶

spective access to pain's badness itself. But what would the latter involve, or even mean, on the objectivist conception of practical reason defended by de Lazari-Radek and Singer? On such a view, pain's badness is presumably an a priori, necessary (even nonnatural) truth. It is rather obscure, at best, how we could know such truths by introspection. In any event, there is also a powerful evolutionary explanation of why, whenever we experience pain, we are also strongly disposed to believe that this state is bad. Can we be confident that, when we form this belief, we really do so through introspection, free from any evolutionary influence?

36. My argument here has focused on de Lazari-Radek and Singer's attempt to use evolutionary considerations to defend impartial conceptions of ethics. But the point is more general. For it seems that any attempt to employ evolutionary debunking arguments to defend substantive normative conclusions would face a similar dilemma.