



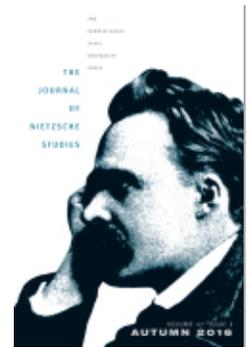
PROJECT MUSE®

“Why?” Gets No Answer: Paul Katsafanas’s *Agency and the Foundations of Ethics*

Jorah Dannenberg

The Journal of Nietzsche Studies, Volume 47, Issue 3, Autumn 2016, pp. 418-434
(Article)

Published by Penn State University Press



➔ For additional information about this article

<https://muse.jhu.edu/article/640469>

"Why?" Gets No Answer: Paul Katsafanas's *Agency and the Foundations of Ethics*

JORAH DANNENBERG

Abstract: In this review, I consider Paul Katsafanas's attempt to provide a constitutivist defense of ethics, informed by his rich and original reconstruction of Nietzsche's theory of agency. In particular, I focus on the ambition to combat nihilism (conceived as a special brand of ethical skepticism), by offering a vindication of the authority of ethical values. I offer some reasons to question the viability of this strategy in general, as well as some considerations concerning the dispute between Nietzsche and Schopenhauer, which lead me to wonder about attributing such a strategy to Nietzsche in particular. Rather than reading Nietzsche as sharing the constitutivist's defining ambition, I suggest that contemporary ethical theory may have more to learn from Nietzsche's diagnosis of what the constitutivist is trying to do. Nietzsche's understanding of nihilism suggests that seeking irrefutable foundations for our values may itself be a symptom of, rather than the cure for, the kind of ethical skepticism that plagues us.

Keywords: Schopenhauer, moral skepticism, nihilism, constitutivism, agency

Introduction

Agency and the Foundations of Ethics is an ambitious, engaging, and challenging book.¹ The foundational problem of ethics, Paul Katsafanas tells us at the outset, is providing a justification of morality's authority, one that can fend off skepticism. Constitutivism undertakes to do just that, by giving an account of the nature of action in terms of some constitutive aim, which will at once vindicate the authority and illuminate (at least some of) the substance of practical normativity. Such a strategy is, Katsafanas argues, uniquely poised to succeed in providing a foundation for ethics, but it has so far failed to deliver. We thus need a new and improved version, and that

is precisely what Katsafanas offers: a version of constitutivism informed by the theory of action he mines from Nietzsche, engineered to succeed where other constitutivist theories fail. Nietzschean constitutivism, Katsafanas argues, can do two things at once: provide us with a clearer and more systematic understanding of certain core elements in Nietzsche's ethical thought, and deliver on constitutivism's inherent promise, to surmount the problems of contemporary ethical theory.

Katsafanas's book thus represents *both* an original contribution to contemporary ethics, and an insightful work of Nietzsche scholarship. I should say right up front that what little "expertise" I can bring to bear concerns only the former. I have studied some of the ethical theory of the late twentieth and early twenty-first centuries, and in particular the corner of late twentieth-century ethical theory where constitutivism first emerged, then became a growth industry. I am, alas, highly unqualified to engage with this book as a Nietzsche scholar would. I confess up front what I think will be painfully obvious soon enough: I am at best half the philosopher I would need to be in order to take the full measure of this impressive book.

With respect to that part that I do not feel totally unqualified to consider, I suppose I might summarize my reaction this way: I have my doubts, and I want to keep them. One thing I do *not* doubt, however, is whether Katsafanas is right to argue that constitutivism is uniquely well positioned to provide the sort of foundational defense of ethics that is aimed at thwarting skepticism. I dare say constitutivism *evolved* for precisely that purpose.² But where Katsafanas sees a reason to try to fashion a new version of constitutivism, I find myself inclining in the opposite direction: I see a reason to be that much more wary of the constitutivist project. I tend, at least some of the time, to think that we contemporary ethical theorists too readily take it for granted that we know what our real problems are, and what it would mean for us to solve them. I am, in other words, suspicious that the aims and ambitions that tend to organize and animate contemporary ethics may be more than a little misguided and confused. This is especially so when it comes to the aim that Katsafanas and I agree drives the constitutivists: providing ourselves with a convincing justification of morality's authority, one that can relieve us of all our skeptical doubts.

To the extent that I find myself increasingly drawn to hearing more of what Nietzsche has to say, this is largely because what little of Nietzsche I think I know has nursed my sense of discontent with contemporary ethics and its defining problems. That, I should again stress, reveals next to

nothing about the scholarly merits of Katsafanas's reading of Nietzsche. But I suppose if Nietzsche really *is* best understood as providing one sort of solution to the problems of ethics as we have come to think of them, I am perhaps a bit less interested in hearing from Nietzsche than I thought.

In sum, my reservations about Katsafanas's book are very much those of a somewhat dissatisfied ethicist, one who doubts whether we in contemporary ethics really understand our own problems, and who doubts even more that we know what it would mean for us to solve them. I want to use this comment to explore, a bit further, just what we should think the problems of ethics *really* are, and in what sense we might or might not be able to see Nietzschean constitutivism as offering us an attractive solution.

Skepticism and Nihilism in Contemporary Ethics

I think it is safe to say that the problem of justifying the authority that morality purports to have is more or less our direct inheritance from Kant. As Katsafanas sees it, however, it is a problem that has gotten harder rather than easier since Kant's attempt to solve it himself. It has gotten harder because we have come to see ourselves as having more reason for being skeptical about morality's authority rather than less.

Katsafanas's discussion begins by identifying three sorts of skeptical challenges that any foundational defense of ethics must overcome: one epistemological, one metaphysical, and one practical. Epistemically, we lack confidence in the truth of our particular ethical beliefs, our views about what is of value, or what is to be done, or our ideas about how to live. Metaphysically, we doubt whether there could even be values, or facts about what is right or wrong, or anything like bona fide normativity at all. We struggle to see how these sorts of things can be integrated into a more or less naturalistic understanding of the world, and of our own psychology. Practically, we worry that the sort of "grip" on us that we take our values to have may be illusory or precarious or both—that nothing about our values can, on reflection, sustain our allegiance to them. We fear that our values, and so our lives insofar as they are shaped by our values, will be revealed as ultimately unsupported, and unsupported.

While versions of these skeptical challenges are familiar, Katsafanas deftly shows how Nietzschean ideas can be marshaled in order to put each one in an especially pointed way. Concerning morality's epistemology,

Nietzsche draws our attention to the fact that morality has a history, one which can provide a powerful explanation of how our current system of moral beliefs developed, which undermines our claim to know anything about what we *really* must or must not do, what we should value, or how we ought to live. On the metaphysical front, Nietzsche's methodological naturalism, and in particular his knack for unmasking the *actual* workings of the human psyche, make it seem preposterous that there might be the sorts of entities, powers, faculties, or processes that moral theories tend to posit: moral reasons, Kantian wills, perception-like faculties of moral intuition, agent-causation, and on and on—these come to seem like so much superstition in the wake of Nietzsche's gimlet-eyed investigations. Finally, there is the third challenge, which Katsafanas labels the "practical challenge." I am going to focus on it, for I think that from the point of view of contemporary ethics it is clearly the most novel, arguably the most important, and by far the least acknowledged and understood of the challenges Katsafanas considers.

Katsafanas introduces the practical challenge as that of explaining "how and why morality has its grip on us."³ He rightly rejects the question of "motivational judgment internalism" as an adequate way of characterizing the problem, suggesting instead that Nietzsche's discussion of nihilism can help make out the real challenge in an especially poignant way. The threat of nihilism, as I understand it, is the threat that values might come to seem arbitrary to us and, thus, that we might come to feel that nothing is really worth valuing at all. I take it this is not just a concern about *some* values, though it might start out that way; it is in fact a fear about values *generally*. The threat of nihilism is the fear of losing confidence not only in *our own* values, but in the very *idea* of value—the idea that anything is worth doing, or caring about, or being inspired by, or paying respect to. For the nihilist, and here I follow Katsafanas in quoting from Nietzsche, "life is no longer worthwhile, all is the same, all is in vain."⁴ And again, "The goal is lacking; 'why?' finds no answer."⁵

Katsafanas at one point characterizes nihilism as "the belief that no values are justified."⁶ But I find this characterization doubly misleading. First, though we would no doubt be correct to attribute that belief to the nihilist, talking in terms of belief makes it sound like nihilism is a *doctrine*. But I hear in the remarks of Nietzsche that Katsafanas approvingly quotes something rather more like the characterization of a *condition*—one of despair, hopelessness, boredom, and alienation. Second, while that sort of condition

may be the result of having gone looking for a justification and not finding it, this strikes me as inessential, or perhaps as representing a more advanced stage of the condition itself. At any rate, we should not want to rule that out by definition. The original worry—the first kernel of nihilistic doubt—seems to me not a matter of wondering whether our values are *justified*, but rather *justifying*—not a matter of whether *we* can provide support for *them*, but whether *they* can provide support for *us*. Perhaps one naturally and inevitably goes looking for something further to support one's values once one begins to sense that living in accordance with them is not enough to ensure that one's life is worthwhile. But the sense that one's life or one's values *require* some further support can, it seems to me, be just as readily construed as itself a symptom of the very condition we mean to diagnose.

I think Katsafanas is entirely right to claim that this sort of nihilism represents a distinctive challenge, and one that is ever-present, if not overtly so, throughout contemporary ethics. As he points out, more familiar are attempts to rebut or condemn the amoralist or the egoist. But neither Thrasmachus, nor the Hobbesian Foole, nor the “ideally coherent Caligula,” is a nihilist. Each, after all, has goals and purposes that clearly, for him, “inspire faith.” If ethical theorists are worried about nihilism—and I wholeheartedly agree that they should be—this is not ordinarily apparent in how they present or understand themselves. More work needs to be done in order to understand the distinctive challenge or threat that nihilism poses, and to understand its relationship to the practice of ethical theory. Katsafanas has done us a service merely by calling attention to this.

The Nihilistic Condition

When David Enoch asks what reason he has to perform any actions at all, he is attempting to make a philosophical point. One can indeed refute Enoch, as Katsafanas and other constitutivists do, by explaining why the question “what reason have I to do anything?” is not well formed—that, strictly speaking, the question asks something incoherent. But it does not follow, and it clearly is not true, that, as Katsafanas puts it, “the question whether there is a reason to perform any actions at all is *moot*.”⁷

The question is not moot, because while we may not be able to provide a content for the question, nor can we offer one kind of answer to it, we can certainly understand something about the person who asks it. As

Bernard Williams pointed out, someone who sincerely asks “why should I do anything?” might naturally be taken to express despair and hopelessness. In fact, Williams seems more or less to agree with the constitutivist’s semantic point, when for instance he writes, “it is very unclear that we can give the man who asks [why should I do anything] a reason—that, starting from so far down, we could *argue* him into caring about something. . . . What he needs is help, or hope, not reasonings.”⁸

The man who asks, *sincerely*, “why should I do anything?” seems to me rather close to exhibiting the same condition as Nietzsche’s nihilist. If that is right, I am inclined to think at least the following about the nihilist and how he might or might not be helped. First, someone who has in fact become a nihilist does not appear to be in much of a position to help himself—not even in principle. That is, it is very unlikely that anything he can say to or for himself will cure him of his nihilism. Second, it seems to follow that if *we* are in any sort of position to help him find relief, it must be that we are in a different shape than he is; in particular, we are not (yet) as desperate and hopeless as he, and that is why we may be able to help him to find hope in a way that he could not, even in principle, hope to help himself. Third, pointing out to him that his question makes no sense will be of absolutely *no* help to him; we surely cannot show him how to hope by explaining to him why it is that, as Luca Ferrero elegantly puts it “agency is closed under the operation of reflective rational assessment.”⁹ Nor can we help him by pointing out to him that, by asking for help in the first place, he is already “committed” to there being reasons, and so he has already in effect presupposed *some* kind of answer to his own question. Fourth, whatever genuine help or hope *he* might receive, it will not and cannot come from our providing him with a justification for the authority of some claim or claims made upon him, concerning how or what he *ought* to be doing or feeling or valuing. Help might indeed come from being told what he must do; but part of his very predicament is that if he is even able to follow such orders, it will be in a diminished and attenuated way. Help or hope must come *first*, and then he may be in a position to ask after justifications.

I think Katsafanas and other constitutivists will want to disagree with me on at least one, perhaps some, and maybe all of these claims about nihilism. I think they think that constitutivism, *the theory*, can thwart nihilism, *the condition*, by securing for the nihilist—and us, insofar as we fear becoming more like him—a justification for the authority of some value or values. I see evidence for this ambition of thwarting nihilism in this way, which

I am inclined to think is deeply mistaken, throughout the constitutivist canon. But I will here focus on one piece of Katsafanas's argument that I find especially telling in this regard: Katsafanas's reconstruction of where Nietzsche and Schopenhauer part company, and his attempt to show, through something like a transcendental argument, that Nietzsche is right and Schopenhauer wrong.

The Basics of Nietzschean Constitutivism

To get to that argument, I will first need to put before us considerably more of the substance of Katsafanas's view. Katsafanas's general defense of the constitutivist strategy, together with his critique of extant versions that comprise the first half of the book, pave the way for Katsafanas's positive proposal: Nietzschean constitutivism, which posits two constitutive aims, *activity* and *power*.

Activity first. Katsafanas accepts, from a tradition with roots in Locke and Kant, the basic thought that in *acting*, an agent may be more or less *active*. This is not, he stresses, the same distinction sometimes thought to be the basic one drawn in action theory, between those events that are actions, and those that are not. Rather, it is a way of grading or evaluating those things we have already granted *are* actions, as more or less paradigmatic or exemplary instances of their kind. The thought, insofar as I understand it, is that some of the things I do may perfectly well qualify as actions, if action is contrasted with "mere" behavior, but *I* may nevertheless be less than (fully) active in their production. In other words, we need a way of making sense of the idea that there are things that I undeniably, in one sense at least, count as *doing*, despite the fact that I, in some other sense, am merely along for the ride when they are done. Examples, I gather, include certain sorts of actions commonly thought akratic—for instance, having that third drink despite my judgment that I had better not. Or, perhaps, an impulsive and insulting outburst that is due to neither a brute compulsion like Tourette's, nor to a fully calculated or deliberate intention to cut someone down. In such cases, we might think that while I most definitely do *act*, *I* am nevertheless more spectator than agent with respect to what I do.

To draw any such distinction, Katsafanas argues, we must accept—along with Locke and Kant—that our motives incline without necessitating—that we have some ability to "step back" from a motive and deliberate, in a way

that actually exerts some influence on how we go on to behave. But the tradition tends to go further, attributing to us the power to suspend entirely the influence of our motives. For the tradition, a human agent is fully active just in case, having stepped back from *all* of her motives and evaluated them independently of *any* influence upon her, *she* determines her action unimpeded by any alien influence. Such a picture is, Katsafanas argues, philosophically untenable, false to our experience, and refuted by empirical psychology. We do not have the ability to deliberate and choose from a perspective entirely outside our various feelings, drives, moods, desires, and so forth. There simply is no such unsaturated perspective for us to occupy—no “view from nowhere” to which we can retreat, no purely rational self to do the retreating.¹⁰

If the active-passive distinction in action is to be maintained, it must therefore be redrawn in order to incorporate the basic psychological facts. This Katsafanas does by proposing a criterion of activity in terms of what he calls (borrowing from Nozick) *equilibrium*: roughly, an agent is *active* with respect to some *action* just in case she approves of the action, and further knowledge about the motivational etiology of the action would not undermine her approval. In other words, I need not—as the tradition had it—be fully, or even minimally aware of nor in control of what moves me, in order for some action performed by me to count as, in the fullest sense, “mine.” What makes an action *mine* is, rather, that I approve of what I do, and that I am disposed to continue to approve, were I to have all the facts about what I was up to put before me.

Katsafanas argues that being *active* so conceived is a *constitutive aim* of action. This means that insofar as we act, we *must* undertake to act in ways that are in the fullest sense our own, which in turn requires us to take our own approval—both in fact and in disposition—as normative. So, activity-as-equilibrium establishes the first standard that every agent must strive to meet, and against which every agent must measure herself in acting. This standard would, however, appear to make the content of normativity depend on an agent’s particular, idiosyncratic values. For it is fundamentally in light of these values that she will, as a matter of fact, approve or disapprove of her action. I may, for instance, be the sort of person with a system of values that leads me to approve of my actions when they inflict harm on others. You may be the sort of person whose values lead you to approve of your actions when they exhibit compassion. Nothing in the constitutive aim of activity could adjudicate between these systems of values,

each of which would presumably yield radically different results about how each of us ought to act.

But in fact, the lack of determinate content and the related failure to adjudicate between conflicting sets of values is not, I believe, the real problem for Katsafanas. Because this story bottoms out in facts about a particular agent's values, Katsafanas thinks an agent would be *mistaken* if she understood her own values as having a legitimate claim to *authority* over her. I think what he must have in mind is something like this: confronted with a situation in which it would be difficult to muster an action of which she could approve, an agent would be under no normative pressure whatsoever to stick with her values; she could, so far as the source of normativity is concerned, simply dispense with her old, inconvenient values and adopt new, more convenient ones. Any stability her system of values exhibited would thus be due to inertia, or luck. What Katsafanas wants is something rather more than that: an authoritative standard, one that can tell her when and whether she *must* stay devoted to her values. Enter *power*, action's second constitutive aim. In every action, Katsafanas argues, an agent necessarily aims to express her power.

Katsafanas is here making a claim about the very essence of willing, which he presents through an explication of Nietzsche's concept of *will to power*. Unpacking the idea first requires elaborating the structure of *drives*—a particular kind of motivational factor that, unlike the basic elements in many more familiar models of our practical psychology, are not in the first place a matter of an agent's being directed toward some determinate goal or end. Drives may, and often do, orient agents toward ends or goals, but this is a secondary consequence of the drive's more fundamental aim, which is simply its own *expression*. My aggressive drive, for instance, might move me to belittle someone. Hurting that person's feelings indeed becomes my goal, but it is not, as it were, that for the sake of which I act. Rather, the aggressive drive's fundamental aim is simply its own expression. The vulnerability of another merely happens to provide opportunity.

Like aggression, our aiming at power is not, Katsafanas urges, to be understood as aiming at some determinate *end* or *goal* in action—he is not claiming that we necessarily act in order to attain power, or even in order to exercise it (at least where “in order to” is understood in its usual sense). Power is expressed through the encountering and overcoming of resistance, in the pursuit of whatever else our goals or ends happen to be. But power is not to be understood as *itself* a drive, to be placed alongside our other

particular drives; that would be a kind of category mistake. Rather, power is something like the very *form of a drive*. For any particular drive to successfully express itself *just is* for it to encounter and overcome resistance. Thus, *will to power* is a formal feature of how a drive—any drive—structures the pursuit of its (happenstance) ends or goals. Any drive aims to express itself *powerfully*. The idea is that whatever else it is that we aim to do, or will, or bring about through acting, insofar as we are acting on drives we are necessarily aiming to express them, and thus necessarily aiming to realize our ends or goals in a powerful way.

To say that *all* action aims at power, it thus turns out, just is to say that *all* action is motivated by drives—all action is aimed at the successful expression of some drive or other, and so *necessarily* has the higher order aim of every drive, to encounter and overcome resistance. And indeed, this is what Katsafanas, following Nietzsche, argues: our psychology is a drive psychology. Whenever human beings act, they act on drives, and so every human action has the constitutive aim of power.

Power can thus provide a kind of authoritative standard, Katsafanas argues, against which the otherwise arbitrary values that ground the evaluation essential to achieving action's first aim, activity, can be assessed. Values that tend to ground approval of actions that are not or cannot be willed *powerfully*, will be defective with respect to this higher-order aim. The aim of activity underdetermines what we have reason to do, insofar as it is silent on the question of what we should value. The second aim, power, breaks that silence, by telling us to value only those things that enable us to will power. The overall result is normative pressure to adopt a system of values and act in ways that will elicit approval in light of that system, in a way that allows an agent to fully realize *both aims*—to be at once active and powerful.

Nietzsche vs. Schopenhauer

With all that in mind, let me return to Katsafanas's reconstruction of Nietzsche and Schopenhauer. As Katsafanas reads them (and who am I to disagree?) Nietzsche and Schopenhauer are in close to perfect agreement on at least one fundamental point: that "all actions manifest a ceaseless, indeterminate striving. This striving is not directed toward any particular end, but simply toward activity."¹¹ But where Schopenhauer sees something terrible in this, Nietzsche evidently finds something worthy of affirming. Why,

according to Nietzsche, are we to affirm ourselves, insofar as we recognize ourselves as ineluctably driven in this way? Why not, as Schopenhauer suggests, lament this fact about ourselves, and do whatever we can to try to relieve ourselves of the burden it places upon us?

Katsafanas seems to think that if Nietzsche cannot give a certain kind of argument that demonstrates why Schopenhauer is mistaken in seeking relief in the form of “denial” or “self-suppression,” then he may be in big trouble. What Katsafanas thinks Nietzsche needs is thus a legitimation of will-to-power’s authority, a showing that will-to-power is entitled to command our embrace: “If an agent can coherently regret the presence of inescapable aims and therefore seek the elimination of action, this might undercut the alleged normative authority of action’s constitutive aim.”¹² Katsafanas undertakes to give just such an argument on Nietzsche’s behalf:

1. If the agent performs an action A-ing, she is committed to agential activity. That is, in A-ing, she is committed to approving of her A-ing, and to having this approval be stable given further facts about A-ing’s etiology.
2. The etiology of every action includes will to power.
3. Thus, in order for the agent to be active, the agent’s approval must be stable given further facts about the way in which will to power motivates her.
4. In this sense, the agent must approve of will to power as a motivating force.¹³

I see some problems with the argument,¹⁴ but rather than rehearse them, let me instead pose a question: What sort of effect is such an argument expected to have in the face of Schopenhauer’s pessimism, assuming it is sincere? That is, what exactly is such an argument supposed to do *to*, or *for* Schopenhauer, or anyone already somewhat persuaded by him, in the direction of getting him to affirm and embrace what he presently denies and tries to stifle?

I find the answer to that question especially opaque, in light of the nature of Schopenhauer’s own position, which Katsafanas helpfully reconstructs for me in a rather detailed footnote.¹⁵ Schopenhauer evidently recognizes that it would be absurd to *try* to stop oneself from willing—in the very trying, one would be willing, and so one would have failed. It thus seems that Schopenhauer cannot be understood as trying to offer us even a *pro tanto*

reason for doing *that*. What he can do, and apparently does, is offer some ideas about what will happen to us if we start to pay more attention to the ubiquity and inevitability of suffering, and the conditions that, according to him, necessarily engender it. Such sustained attention will, he tells us, ultimately lead to the will's withering away, and with it our suffering, though again not through any overt act (or series of acts) that would constitute anything *willed*.

Schopenhauer, it seems to me, must be preaching a kind of conversion—trying to show us a *there* to which we cannot get from *here*, at least not by following the routes laid down before us by practical reason. This is what I make of his instructing us in how, if we want our suffering to go away, we might do well to pay more of a certain kind of attention to its ubiquity. He is offering us a prescription in the form of a not-altogether-intentional strategy for finding relief. We will be moved to follow his advice, insofar as we find his diagnosis of the human condition convincing, and his claim to have found an antidote for it credible.

If that is in the neighborhood of right, then it seems clear to me that Schopenhauer needs nothing like a legitimation of authority. When the doctor pulls out the big, ugly syringe and says, "You must have a shot," the question "Do I *really* have to?" is not a question about his normative entitlement to issue you valid commands. It is rather responded to with something like, "Yes, it is the only way." And then of course one may still wonder, "Is it?" And then one might want a second opinion.

Nietzsche, it seems to me, is in a position exactly symmetrical to Schopenhauer: he has his own diagnosis, and his own purported cure. From what I gather, he wants to convince us that an immense amount of suffering can be borne, provided one comes to have the sense that it is *for* something. To the extent that we will tend to experience our suffering as meaningless, something to be relieved, this is due to our alienation from our nature as creatures who ineluctably strive to express our power. In large part, that alienation is the result of our having embraced a system of values—moral values—that pit us against our selves. To cure ourselves of what ails us, we should indeed embrace will to power, and reject old values insofar as they are incompatible with it, Nietzsche tells us.

But again, it seems to me that a claim to authority would here be beside the point; I find it very odd to put one in Nietzsche's mouth. The closest I can come is in thinking of Nietzsche's point as something like the following: we must choose to *make* power our authority, by embracing it. But,

just as in Schopenhauer, the force of this “must” is peculiar, and it surely cannot be the “must” of a rational norm. Relatedly, grasping the sense in which embracing power is something we are in any position to *do* seems to me to require some delicacy. After all, if Nietzsche is right and our wills are already sickly, then we are likely to find that we cannot so easily take his medicine, even if we want to.

I think that for Katsafanas, unless we can settle the dispute between Schopenhauer and Nietzsche with an argument that establishes that we are already “committed” to treating will to power as an authority, we are left with something here that is merely a matter of taste. It is as though we might just have to say: “Some go in for a resolute sense of purpose that makes their immense suffering bearable; others go in for withering away of willing, and with it suffering itself. Who is to say who is *right*?” In one sense, I find myself agreeing that to leave it at this would do violence to the profundity of their disagreement. But I worry that no less violence is done by seeing this as a question of legitimating the claim to authority made on us by our nature. My thoughts return to Williams’s point: if Nietzsche thinks he has something to offer Schopenhauer, and to us insofar as we are like him or likely to be moved by him, it seems to be something more like help, or hope, and not arguments. “You do not have to starve yourself and your desires, or otherwise let your will wither away,” we might imagine Nietzsche saying to us, “for your problem is not that you suffer, but that your suffering cannot mean anything for you. I can show you how you can make it mean something, by showing you what it would mean for you to embrace your nature as a willer.”

The Opposite of Nihilism

Katsafanas has not yet convinced me to embrace that part of myself that aims, ineluctably, at encountering and overcoming resistance. But one thing that Katsafanas has convinced me of is something he probably did not quite intend: that contemporary ethical theory needs to pay *much* more attention to nihilism, and the unique sorts of problems that it presents.

One thing that would help, I think, is if we had a concept for the condition that is the *opposite* of nihilism: the condition a person is in when she and her life stand in the right sort of relation to her ideals and her values, so that they are genuinely her own, and are capable of nourishing

her and propelling her forward. Such a concept would have to be one that allowed us to raise the question, without begging it, of whether a person could genuinely be in that condition, without necessarily being able to “justify” the claims that her values make upon her to any other rational agent as such.

If contemporary ethical theory already has such a concept, I am not aware of it. Bernard Williams appears to have tried, at least at one point, to reclaim the concept of *justification* itself for something like this very role.¹⁶ But that effort, in retrospect, proved a mistake. Williams could not cancel the rationalist connotations of the notion of justification—in particular, the idea that for an agent to be justified is for her to be flush with the legal tender of rationality, universally valid reasons, with which any justificatory debt, public or private, can be paid. His effort thus bred more confusion than anything else. Other words that he could have used, and sometimes did try to use, to capture the particular way in which a person’s values, projects, and principles can, if she is lucky, fund her life—words like purpose, meaning, or ground—invite other, equally problematic misunderstandings in contemporary ethics.

This is highly speculative, but I suspect that if we had a concept like the one Williams tried to make justification into, it would especially help contemporary ethicists like me to get a handle on what Nietzsche might have to say to us. It might also help us to understand how constitutivism, and especially Nietzschean constitutivism, could hope to be true. If there is any sense at all to the idea that one can derive a justification for one’s values, principles, or ideals from the facts about one’s nature as an agent, it seems to me that this cannot be understood on the model of deriving the conclusion of an argument from its premises. To steal an idea from Stanley Cavell, it must be rather more like the way one derives pleasure from playing the piano. One thus cannot expect an answer to “why must I embrace and identify with that part of me that wills power?” if one is not already doing so. One must be helped to embrace that part of oneself, and then—maybe—one will derive the why.

In this respect, I think Nietzschean constitutivism must be understood to be in a rather different position than Korsgaardian constitutivism. Korsgaard, it seems to me, is in the rather advantageous position of selling us an image of ourselves and our nature that most of us are *already* embracing, insofar as most of us do tend to identify with the values and ideals of enlightenment morality. She can, it seems to me, agree with Williams

in thinking that a person who has begun seriously to ask, “Why should I do anything?” should be offered help or hope, not reasoned with. But her diagnosis will be that, insofar as that question comes to seem serious, it is because such a person has begun to lose hold of her reason. Help thus comes in being reminded of what one already knows about oneself, and what this entails. Provided she is willing to try, every agent is assured in advance that she can succeed in “making something” of herself, for the constitutive principles of autonomy and efficacy just are the failsafe recipe for doing so. For Korsgaard, a satisfying answer to “Why?” is the birthright of every rational agent, and she promises to help you find your why by helping you discover what follows from the fact that you are more or less what you would like to believe that you are.

Katsafanas’s Nietzsche, it seems to me, must want to nurse the very doubt that Korsgaard aims to quiet. His recommendation is that we undertake to embrace a part of ourselves that, insofar as we have come to identify with enlightenment values, is bound to seem at least somewhat alien, regrettable, or even shameful to us. It will continue to seem that way even *if* we acknowledge it as an “inescapable” part of who we are. Importantly, he *cannot* offer any assurances: a fully meaningful or worthwhile life is not within the reach of everyone *as such*, and there can be nothing to guarantee in advance that one will succeed in making something worthwhile of oneself by following his prescription. Essential to Nietzsche’s view seems to be that, despite trying with all their might, many people will find themselves left without any satisfying answer to “Why?”

So in effect, Korsgaard says to me: If you are worried that “Why?” will find no answer, do not worry. Instead, double-down on your enlightenment ideals, and then your values are guaranteed to support you and never desert you. Nietzsche says, if you are even a little worried about nihilism, you are probably right to be: that is precisely where you are headed. It may not be too late to right the ship, but you had better be prepared to try something fundamentally different: learn to embrace a part of yourself that now seems alien. If you do, *maybe*, through some combination of effort and luck, your life will come to really mean something for you. But there are no guarantees.

Comparing the two at face value, it is pretty clear which is the more inviting wager. If I am to opt for Nietzsche’s way, it seems to me I must at least have come to be convinced that Korsgaard’s is too good to be true, and thus a lie. I *am* more or less convinced of that, and so not prepared

to double-down on enlightenment morality with Korsgaard and Kant. But Katsafanas's Nietzschean proposition strikes me as a very risky one, and I am not yet prepared to make that bet either. That brings me back to what I said at the outset: I have my doubts, and—for now at least—I want to keep them.

Stanford University
jorahd@stanford.edu

NOTES

1. Paul Katsafanas, *Agency and the Foundations of Ethics: Nietzschean Constitutivism* (Oxford: Oxford University Press, 2013).

2. Christine Korsgaard marks the emergence of constitutivism as a distinctive option in contemporary ethics; see Christine Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996). But many of the essential pieces of the constitutivist program show up first in Thomas Nagel's *The Possibility of Altruism* (Princeton: Princeton University Press, 1970). See esp. chap. 4 and the discussion of what Nagel calls *interpretation*. There are important differences in the two authors' ambitions vis-à-vis skepticism. One should compare the discussions of skepticism from the final chapter of Nagel's book and the opening chapter of Korsgaard's. In effect, Korsgaard picks up a challenge laid down by Nagel, to "raise the cost of skepticism further by pushing the roots of moral motivation still deeper" (Nagel, *Possibility of Altruism*, 143).

3. Katsafanas, *Agency and the Foundations of Ethics*, 19.

4. Z IV: "The Greeting" 11, quoted in Katsafanas, *Agency and the Foundations of Ethics*, 23.

5. KSA 12:9[35], quoted in Katsafanas, *Agency and the Foundations of Ethics*, 22.

6. Katsafanas, *Agency and the Foundations of Ethics*, 23.

7. Katsafanas, *Agency and the Foundations of Ethics*, 52, my emphasis.

8. Bernard Williams, *Morality: An Introduction to Ethics* (Cambridge: Cambridge University Press, 1972), 3.

9. Luca Ferrero, "Constitutivism and Inescapable Agency," in *Oxford Studies in Metaethics*, vol. 4, ed. R. Shafer-Landau (New York: Oxford University Press, 2009), 303–33, 308.

10. I am generally sympathetic to this claim and defend a similar one in "Promising Ourselves, Promising Others," *Journal of Ethics* 19.2 (2015): 159–38.

11. Katsafanas, *Agency and the Foundations of Ethics*, 205.

12. Katsafanas, *Agency and the Foundations of Ethics*, 206.

13. Katsafanas, *Agency and the Foundations of Ethics*, 207.

14. Briefly and inadequately: it seems to me that the first premise equivocates on the notion of commitment, in a way that begs the question against Schopenhauer. Granting that Katsafanas is right that activity is a constitutive aim of action, then Schopenhauer is indeed “committed” to acting in ways of which he approves, but in just the following sense: he is committed to assessing his actions by that standard, *insofar as he acts*. He is, however, not committed to acting—precisely his question is whether to be committed to that. But then, I do not see how appealing to anything *further* about the nature of action, including its second, higher-order constitutive aim, could be thought persuasive for him. Insofar as inaction is a live option for him (albeit not an option accessible via the exercise of practical reason, as he admits), I do not see how the argument can hope to work.

15. Katsafanas, *Agency and the Foundations of Ethics*, 205–6 n. 45.

16. See especially Bernard Williams, “Moral Luck,” in *Moral Luck*, ed. Bernard Williams (Cambridge: Cambridge University Press, 1982), 20–39.