

**Paul Katsafanas**

*Agency and the Foundations of Ethics: Nietzschean Constitutivism.*

Oxford University Press 2013.

288 pages

\$75.00 (Hardcover ISBN 9780199645077)

In *Agency and the Foundations of Ethics*, Paul Katsafanas seeks to provide a general account of normative judgments. The book is ambitious, perhaps too much so, addressing central questions in metaethics, action theory, and normative ethics, in addition to arguing for an interpretation of Nietzsche's views on ethics. The book is to be highly commended for being able to bring together various discussions that often run independently of each other. His attempt to combine constitutivism, often thought of as a broadly Kantian view, with Nietzsche's ethical views is creative and sure to be controversial among interpreters of Nietzsche. Though, as I am about to explain, I have doubts regarding the conclusions he reaches, Katsafanas gives us lots to think about and his choice of what to focus on and overall approach is well conceived.

Katsafanas addresses central questions in metaethics by developing an account of constitutivism where "the authority of certain normative claims arises from the bare fact that we are agents" (1). The general idea is that action has a constitutive aim and "we could derive normative conclusions from that constitutive aim, and thereby show how claims about universal reasons for action might be justified" (68). On Katsafanas's Nietzschean account, there are two constitutive aims of action: agential activity and the will to power. Agential activity is understood by Katsafanas in broadly compatibilist terms: an agent is active if the agent approves of her action and further knowledge of the motives behind the action would not undermine her approval (111). Katsafanas's account of the will to power claims that "to will power is perpetually to seek to encounter and overcome resistance in the pursuit of some end" (159). Katsafanas addresses questions in action theory and practical reason by appealing to the account of agential activity together with a Nietzschean theory of motivation where all human actions are motivated by drives. Finally, he addresses questions in normative ethics by developing the notion that the will to power is constitutive of action and that other values can be endorsed or rejected in terms of the will to power.

The book starts with a chapter devoted to three challenges to ethical theory: the epistemological, metaphysical, and practical challenges. Katsafanas argues that constitutivism is better able to deal with each of these challenges than the alternatives. He describes the epistemological challenge as arising "from a simple fact: morality has a history" (8). He draws centrally on Nietzsche's genealogical critique, but also mentions Marx's and Webber's views and evolutionary accounts of the development of morality. While the details of the views vary, Katsafanas thinks that they all undermine confidence in our current evaluations by providing accounts of morality as resulting from biological, social, or economic factors that do not depend on the truth of moral judgments. A successful theory of morality needs to be able to provide a response to these doubts. Katsafanas presents the metaphysical challenge as a set of requirements: that the theory be metaphysically respectable by not appealing to non-natural qualities and that it be based on a model of agency that is not ruled out by our best scientific accounts. According to Katsafanas, the second of these requirements is the harder one to meet – he indeed ends up criticizing the Kantian account for failing to meet it – but he believes that the first requirement can be met by the right naturalistic theory. The last challenge is the practical challenge, which is based on the idea that an adequate theory must be able to explain how morality gets a grip on us so that it "not only tells us

what to do, but purports to outweigh many competing claims about what to do” (19). When it comes to the practical challenge, Katsafanas once again turns to Nietzsche to press the worry that we could stop seeing any value as authoritative: “the whole system of normative judgments might become detached from practical deliberations” (24). Katsafanas sees the challenge as avoiding nihilism, avoiding a point of view from which normative judgments have no authoritative value and thus nothing really matters.

Katsafanas claims that constitutivism is better able to deal with these three challenges than the alternatives. He argues that non-reductive realism (his discussion is focused on Parfit and Scanlon) fails on all three challenges. He argues that Aristotelian theories which attempt to derive facts from the natures of things (he focuses on Hurthouse, Thomson, and Bloomfield) fare better on the epistemological and also perhaps the metaphysical challenge but they have trouble answering the practical challenge. According to Katsafanas, Humean theories, which hold “that normative facts must be explained by some conative state of the agent” (32), do pretty well with the metaphysical and the practical challenges but not so well with the epistemological challenge. He finally considers the Kantian theories and discusses at some length Velleman’s and Korsgaard’s views in Chapters 3 and 4 respectively. On Katsafanas’s view, the Kantian theories, attempting to derive universal normative claims from facts about agency, could work in principle given that they are constitutivist theories in some ways similar to the one that Katsafanas himself is defending. However, he argues that the Kantian views rely on a misguided account of agency, one that is not naturalistically respectable or supported by empirical results. In addition, he claims that the Kantian accounts fail to derive any substantial normative conclusions. Except for the discussion of Velleman and Korsgaard’s Kantian views, which get two dedicated chapters, the discussion of the other three views is somewhat quick; defenders of non-reductive realism, Aristotelian theories, and Humean theories may not find enough to convince them that constitutivism is a superior view. Nonetheless, Katsafanas dedicates Chapter 2 to making the case for and responding to objections to constitutivism. This discussion is well developed and even if he hasn’t provided enough in Chapter 1 to show the other theories lacking, Chapter 2 may sway some to take constitutivism as a serious alternative to the others.

Chapter 2 defends the view that normativity is found in inescapability. This is a bold claim that faces a number of objections, including for example that it attempts to derive an ought from an is, or that it is simply false: inescapability just does not entail normativity. Of particular note in this discussion is his attempt to respond to Enoch’s challenges to constitutivism. Katsafanas is honest and bold in stating the claim that normativity is found in inescapability and in defending it. The view he is defending is perhaps found in other constitutivist theories, like Korsgaard’s and Velleman’s, but it is not directly stated as such. According to Katsafanas, “all that we need, in order for the constitutivist project to work, is the claim that aims in general generate standards of success” (57). One of the strongest objections to constitutivism is focused exactly on that claim and may be derived from G.E. Moore’s critique of evolutionary accounts of ethics. The idea is that even if we can’t avoid having a certain goal or aim, we may still find that aim lacking in value. We may find ourselves unable to escape the constitutive aims of agency, while nonetheless wishing that we could. We may find that there are certain things that we can’t avoid pursuing or aiming at but we may still disapprove of such pursuits. Attempting to respond to this charge, Katsafanas says: “approval of the constitutive aim, or caring about the constitutive aim, would not be normatively relevant. Caring does serve a function in the case of optional, non-constitutive aims, for these aims often weaken once the agent’s approval dissipates. In the case of the constitutive aim, though, approval cannot have this effect: If the aim is constitutive of agency, it will be ineluctable. Approval of the aim might be nice but it is not necessary” (59). Katsafanas seems to think that approval of the constitutive aim of action is

unnecessary because the aim is inescapable but this does not seem quite right. The aim may be inescapable and yet I may disapprove of it. Like Sisyphus, we may find ourselves stuck aiming to push a rock up a hill but we need not love it. I may like or dislike the way practical reason functions; the fact that I can't avoid thinking in such a way is not enough to make me endorse it.

This problem is not unlike the practical challenge that Katsafanas discusses in Chapter 1. In discussing that challenge, Katsafanas says that there must be something that can be said to the skeptic: "we need to offer some explanation of why these rules have a grip on the agents. Absent such an explanation it is hard to see why the rules shouldn't just wither away" (21). The explanation he proposes on behalf of constitutivism is that certain aims are inescapable thus even the skeptic can't avoid them. I am not sure why this explanation would satisfy the skeptic though perhaps the rules wouldn't just wither away even if they lack justification, because they can't. Furthermore, it is unclear how this response could avert nihilism. Perhaps Katsafanas's response to the skeptic is stronger than it may seem at first sight. The constitutivist claim is that there are certain aims that are central to agency that in addition to being inescapable, are what make us agents. Thus perhaps there is an explanation for the rules' grip on us but we can still imagine the skeptic saying: "fine, I am stuck with agency but why should I care about being an agent, why should I bother fulfilling the aim?"

Katsafanas attempts to respond to this line of objection by developing the response he provides to a similar problem raised by Schopenhauer. As mentioned above, for Katsafanas, there are two constitutive aims of action: agential activity and the will to power. Katsafanas argues that Schopenhauer and Nietzsche "agree on the underlying idea that all actions manifest a ceaseless, indeterminate striving" (205). However, Schopenhauer finds this constant striving problematic because it is a constant source of suffering. Nietzsche, on the other hand, "takes it as something that is normative for us" (205). So, "whereas Schopenhauer takes ceaseless striving as a mark of a life's futility, Nietzsche sees in it life's value" (206). Katsafanas endorses the Nietzschean position and he defends it by combining the two constitutive aims in such a way that if someone fails to approve of the constant striving based on the will to power, they are not active and thus not really disapproving. "In short, aiming at activity commits us to approving of will to power ... Merely in virtue of acting, we become committed to *approving* of will to power" (207, emphasis added). This is a clever move on Katsafanas's part, however, in my view, it is sweeping the major concern with constitutivism under the rug. Notice the structure of the argument "we are committed to acting in ways that we can endorse. The only way we can do that, though, is to endorse will to power. So the only way we can *act* is to endorse will to power" (207). So, unless you approve of will to power (the aim that is inescapable), you are not *acting*. But the skeptic and Schopenhauer now join forces and wonder, "why should I care about acting?"

I have a lot more to say about Katsafanas's book than what I am able to say here. The book is thought provoking and worth engaging with. It will be of interest to those interested in metaethics, moral psychology, and action theory in addition to those interested in Nietzsche's views about ethics and on how Nietzsche's views may fit into contemporary discussions.

**Ariela Tubert**

University of Puget Sound