

people on both sides are paraplegics, but . . . if we save people on the right, we will unparalyze them; if we save people on the left, we will not unparalyze them” (453). Suppose we agree with Kamm that in this case we should toss a fair coin to decide whom to save. Does that show that being unparalyzed is an irrelevant good in a life-and-death contest? Here is a case which suggests not. Suppose that, unlike in Kamm’s example, if we assist people on the left side there is a 11 percent chance that we will save their lives and a 1 percent chance that, if we save them, they will become unparalyzed, whereas if we assist people on the right side, there is a 10 percent chance that we will save their lives but a 99 percent chance that if we do so, they will become unparalyzed. In this case, it seems the fact that we stand a much better chance of unparalyzing people on the right side speaks in favor of using a lottery device that gives them a greater chance of being saved. If so, it is not true in general that the good of being unparalyzed is irrelevant in life-and-death contexts.

Despite the ingenuity of Kamm’s arguments it can safely be predicted that few will feel persuaded by them on all of the many questions addressed in this book. Still, it is extremely difficult not to be impressed with her powers of imagination and her moral perceptiveness. Many of the individual chapters brim with crucial insights, decisive counterexamples, and puzzling questions, and together they offer a top-notch deontological exploration into various matters of bioethics as well as into normative ethics in general. While little attempt has been made to give the collection the form of a coherent, well-shaped monograph—a number of arguments, examples, and principles appear in more or less the same form in different chapters of the book—it is still a very useful book. I strongly recommend it to students and scholars of bioethics as well as to philosophers interested in normative ethics in general.

KASPER LIPPERT-RASMUSSEN
University of Aarhus

Katsafanas, Paul. *Agency and the Foundations of Ethics: Nietzschean Constitutivism*. Oxford: Oxford University Press, 2013. Pp. 267. \$75.00 (cloth).

This book defends a version of constitutivism inspired by the moral psychology of Nietzsche. A constitutivist theory is a metanormative theory that purports to establish the objective validity and the content of practical norms on the basis of the constitutive features of agency. At the outset of the book, Paul Katsafanas offers a most helpful general introduction to constitutivism as a kind of metanormative theory and its relation to traditional metaethical views. In chapter 2, he defends constitutivism against the most pressing objections. In chapters 3 and 4, he discusses and criticizes the two most prominent constitutivist accounts, those of David Velleman and Christine Korsgaard. In chapters 5–9, he develops his own version: Nietzschean constitutivism.

In this review I will focus exclusively on the constructive portion of the book. I won’t discuss Katsafanas’s criticisms of Velleman and Korsgaard and his interpretation of the historical Nietzsche (which he discusses at length in the appendix).

Katsafanas focuses on the version of constitutivism in terms of the constitutive aim of intentional action and the standards of success engendered by this aim. According to him, there are actually two inescapable constitutive aims of intentional action: the stable approval of motives and the 'will to power'. First, reflective action aims at securing 'agential activity', the distinctive contribution of full-fledged agents to action. Katsafanas argues against the standard view of agential activity, which takes as paradigmatic of this activity the actions produced by deliberation and reflective choice. Although reflection and deliberation might be able to stop the influence of motives in determining one's actions, there is empirical evidence that reflection cannot stop the influence of motives on deliberation itself (130–31). Reflective actions could thus count as instances of agential activity only if, were the agent to come to know about the unavoidable influence of motives on her deliberations, she would continue to approve of these motives. Therefore, agential activity constitutively aims at securing not only that one acts in accordance with one's choice but also that one's approval of the underlying motives is not undermined by further knowledge of how these motives figure in the etiology of one's action (including the deliberation leading to it; 100–111, 138, 143).

Aiming at stable approval only provides a structural constraint, since it does not offer a substantive standard for the approval of motives. Substantive normative guidance is rather to be found in a second constitutive aim of reflective action: the 'will to power'. All human actions (reflective and unreflective alike) manifest will to power in that they inescapably aim at encountering and overcoming resistance in their performance (176).

Will to power is the form taken by willing for agents with a 'drive psychology' (170). Drives are dispositions that induce affectively charged orientations and dispose the agent to engage in a particular activity rather than to bring about a determinate end. Once activated, drives continuously seek their expression and discharge (166–67). Unlike a desire-motivated action, which usually aims at bringing about an end that would put the action to rest, a drive-motivated activity is directed at the performance of the activity itself (169). The pursuit of particular ends is only an occasion for the manifestation of the drive. It follows that will to power can never be permanently satisfied; it rather involves a 'perpetual striving' (157–58). Katsafanas claims that all human actions are drive motivated. This is shown not only by contemporary empirical psychology (174) but also by Nietzsche's observation that for beings like us there is no state that could provide lasting satisfaction (172). Happiness should not be conceived as a state of lasting satisfaction, which we couldn't possibly achieve. For us, satisfaction is processual: what we aim at is "simply to be active . . . to engage in the process of encountering and overcoming resistances in the pursuit of ends" (175).

Any constitutive aim gives rise to a *pro tanto* reason to fulfill that aim (39). In the case of will to power, the reason is to seek actions that engender the encounter and the overcoming of maximal resistance. This *pro tanto* reason is only one important factor in the determination of the balance of reasons (185). The agent's other values, which are often independent of the will to power, provide additional normative standards (187). When these standards conflict with the will to power, the latter takes priority because it is inescapable and unalterable. As a result, the values in necessary conflict with the will to power or one's capac-

ity to pursue it—such as the disvaluing of suffering and the valuing of certain forms of asceticism (195)—must be altered or rejected (188).

Once the conflict is removed, the *pro tanto* reason to seek maximal resistance is to be balanced with the *pro tanto* reasons generated by other aims. Hence, it might turn out that in particular circumstances there is no all-things-considered reason to best fulfill the will to power. But in the long run the balance of reasons “typically favors actions that best fulfill will to power.” This is for two reasons: (a) the will to power is an ubiquitous drive, and (b) usually this drive does not oppose the operation of other motives, it rather reinforces them by modifying the manner in which they are pursued. For instance, the desire for food rather than being directed at eating the same food on every occasion is modified so as to seek a variety of foods and thereby encounter more resistances in the process of its fulfillment (198–200).

Katsafanas discusses a few substantive implications of Nietzschean constitutivism. He first presents the case study of egalitarianism to show that, contrary to Nietzsche’s own view, Nietzschean constitutivism supports a restricted form of egalitarianism (288ff.). In a less defensive mode, Katsafanas also argues that will to power “urges us to strive for a challenging vision of flourishing” by “enjoining us to engage primarily in creative endeavors” (213).

Katsafanas claims that his constitutivism has “modest” ambitions (241). This might partly explain why his work presents only a few substantive normative implications. But I am troubled by the very limited set of directions that his theory offers to attempt to derive more substantive moral claims. It would be easier to see how the theory is to be applied if the theory required us to maximize resistance. But Katsafanas explicitly claims that the Nietzschean view does not generate such a “crude” injunction (189). What we should rather do is to balance the demand for the maximization of resistance together with the demands generated by other values. Unfortunately, Katsafanas’s theory does not tell us much about this balancing (let alone about the source and validity of the other values).

There are two straightforward ways of balancing the maximization of resistance: first, rejecting values that necessarily conflict with will to power; second, pursuing other values in a resistance-generating manner, everything else being equal (as in the case of food presented above). The last qualification is crucial. If we consider two distinct sets of values, both similarly available to the same agent and where no element is in necessary conflict with will to power, how are we supposed to determine which set is to be preferred? Is it the one that maximizes resistance and minimizes conflict with will to power? At times, Katsafanas appears to suggest so (e.g., 240). If so, however, his theory would inherit many of the problems of the crude injunction to maximize, including the issue of how to measure and compare degrees of resistance across different activities, times, and agents.

Much of Katsafanas’s discussion of the substantive implications is focused on the relation between will to power and single values rather than sets of values. The one notable exception is the discussion of why the will to power does not make murder, oppression, torture, and the like into exemplary actions, given that these kinds of acts might appear to maximize resistance. His solution is to claim that there are other values that “most of us embrace” and that would

conflict with acts of oppression, cruelty, and domination (189). Let me quote in full Katsafanas's conclusion: "We accept the following normative claims: murder is wrong, human life is valuable, gratuitous infliction of suffering upon others is wrong, etc. . . . These normative claims entail that murder is disvaluable. So, on the Nietzschean view, we would have to ask whether there are other acts that produce just as much resistance as murder, but do not violate our evaluate commitments. And, as Nietzsche makes clear, there *are*. Nietzsche's favorite examples of such acts are . . . artistic and intellectual endeavors" (190).

This conclusion is reassuring, but the constitutivist account is silent about the origin of these other commitments and on the issue of why we should stick to them once they are shown to be compatible with will to power. Moreover, we are not given any argument to the effect that the alternative acts generate as much resistance as the most heinous ones. What would we do if it turns out that the alternative acts generate less resistance? Couldn't this trigger a reassessment of some of the other commitments, so as to drop or alter some values in order to engender maximal resistance? And how is the theory supposed to handle those agents who happen not to share these commitments? There is no guarantee that the commitments are shared, as Katsafanas acknowledges when he remarks that they are the commitments that "most of us" (but not all) embrace. Hence, the Nietzschean constitutivism proposed by Katsafanas appears to run out of resources to handle the most radical challenge raised by agents who are moved by the will to power but either have quite different commitments from most of us or are willing to radically revise these commitments under the pressure of the will to power.

Perhaps Katsafanas might still be able to show that murder, oppression, and torture are necessarily ruled out by the structural and constitutive features of agency. Even so, this might not suffice to handle other conflicts, those that are likely to arise among agents who embrace quite different sets of values and commitments, especially given the variety of commitments that we should expect on account of the 'parametric universalism' supported by Nietzschean constitutivism. As Katsafanas points out, the will to power is a universally valid normative standard, but different values are appropriate for different types of agents, depending on their traits and abilities (217–18). I fear that this universal standard does not appear sufficient to handle some basic interpersonal conflicts, many of which would arguably fall under the purview of a standard ethical theory.

How universal is this standard? It applies to all agents who have a will to power. But this drive is only a contingent feature of the psychology of intentional and reflective agents. According to Katsafanas, there is empirical evidence that human beings have such a drive, but "we might have evolved differently and not be motivated by drives," in which case "we would be subject to different normative claims" (241).

This conclusion raises two worries. First, how good is the evidence that all human beings necessarily have a drive psychology? It is undeniable that we have some drives, but Katsafanas's thesis is much stronger: all of our actions are motivated by drives. The empirical evidence cited by Katsafanas amounts simply to the claim that "human beings are most satisfied when engaged in activities that provide them with challenges that are neither too easy nor beyond their capacities" (174). This might be true, but it does not seem sufficient to prove the

strong statement that Katsafanas needs. Couldn't it still be possible that some satisfaction is not processual and at least some actions are not motivated by drives? And even if satisfaction depends in part on encountering resistances, this does not show that we are always motivated to maximize such encounters. Finally, it seems possible that the push for being continuously active might at times come not from our psychology but from the nature of the world, which makes many of our attainments necessarily unstable. Second, even if we grant that we have a drive psychology, should we rest content with an ethical theory that is in principle unable to say how we should interact with agents who, like us, are capable of intentional, reflective, and rational activities but, unlike us, lack the will to power? Shouldn't an ethical theory aspire to be less psychologically parochial?

Some readers might be disappointed to find out that a book that bills itself as being about the 'foundations of ethics', as evinced by the title and the initial presentation of the metanormative ambitions of constitutivism, issues into something much less encompassing. As the book progresses Katsafanas makes clear that his account is not going to deliver all the goods that one might have expected, especially when he insists on the nonfoundational character of his positive view (241). But at times he also appears to overstate the scope and import of the view. For instance, in the closing pages of the work he claims that the standards generated by the two inescapable aims of action amount to a "full ethical theory" (240). He also claims that some of the normative results of his theory can be "often surprising" (242) and at times "deeply incompatible" (241) with our current evaluative framework, but I must confess that I did not get such an impression. In part, this is because there are few substantive claims that the theory actually makes, and in part because I did not find the "ethic of attaining, rather than of attainment" (243)—which is how Katsafanas summarizes the basic substantive lesson of Nietzschean constitutivism—to be especially surprising or deeply incompatible with current ethical views.

Katsafanas's most ambitious claim is that his account answers the "foundational question in ethics" (243) by proving the objective validity of at least one universal normative standard. Central to his constitutivist argument is the inescapability of will to power. Unlike other versions of constitutivism, the inescapability in question is contingent. It is not based on the nature of reflective agency as such but on the fact that our psychology has evolved to include the pervasive and unavoidable operation of the will to power. One might worry that this inescapability is not sufficient to answer the foundational question. How could the inevitability of a contingent psychological mechanism ground the validity of an objective practical norm? Couldn't we simply see ourselves as 'stuck' with this inevitable psychological operation but fail to endorse it and thereby not take ourselves to be normatively bound by its standard of success?

Katsafanas sees the problem, but he argues that we cannot but endorse the operation of will to power. By the very nature of reflective agency we inescapably (and noncontingently) aim at agential activity. If we didn't approve the operation of will to power, which has a pervasive and unavoidable influence on our deliberations, we would fail to be agentially active. For agential activity requires that we would continue to approve of the motives that inevitably influence our actions and deliberations if we were to come to know of this influence. Hence, we cannot but approve of will to power. Or so Katsafanas argues (204–8). His argu-

ment requires not only that we inescapably aim at agential activity but that we inescapably succeed at it. But why should an aim, even an inescapable one, be guaranteed success? True, if we fail to approve of will to power, we would be condemned to the impossibility of being agentially active. This would be a most upsetting outcome, but I am not yet convinced that Katsafanas's Nietzschean constitutivism has refuted the possibility of a deep Schopenhauerian pessimism about the human condition. It is in principle still open to us to repudiate the inevitable operation of the will to power, although at the price of succumbing to the inevitable and systematic failure of our striving to be agentially active.

LUCA FERRERO

University of Wisconsin–Milwaukee

Lebron, Christopher J. *The Color of Our Shame: Race and Justice in Our Time*. New York: Oxford University Press, 2013. Pp. xi+202. \$74.00 (cloth).

In *The Color of Our Shame: Race and Justice in Our Time*, Christopher Lebron addresses the unequal status of blacks compared to whites in US society as a moral matter. Lebron's remedy is to focus neither on racism as a vice in the hearts and minds of individuals nor on institutions as impersonal social structures but, instead, on institutions as permeable to the values and goals of individuals. Lebron's moral approach to racial bias is thus innovative and comprehensive. Chapter-by-chapter attention to how he constructs his arguments—for Lebron's is a well-reasoned approach—is rewarded by an almost nostalgic sense of public morality. According to Lebron, racism, in both the individual and the structural senses, is neither inexorable nor out of the control of the body politic.

Lebron's preface might surprise some. Having grown up poor in Manhattan's Lower East Side, he was admitted, as a graduate from Baruch College, to MIT's program in political science in 2005. His goal was to study John Rawls's international political theory. Lebron's dorm was known as "Chocolate City," because the few blacks in attendance at MIT traditionally lived there. His introduction to the social life of the department was a nonstarter—as the sole urban black, Lebron was ignored by a cohort of preppy whites. The Rawlsian literature about imaginary and remote scenarios soon seemed unimportant. Lebron formed a new goal that became the intellectual motivation for the present book (ix–x): "to redescribe the problem of systemic racial inequality as that of social value—the notion that blacks hold a lower place in the scheme of normative value. This is a scholarly way of saying that racial inequality is basically about the way blacks are not really part of our society in the deep way democracy presupposes, requires, nay demands" (xi). Lebron posits the "disconnect" between stated ideals and a reality of indifference, for example, in the US government response to black victims of Hurricane Katrina, as a reason for shame: "For we consistently undermine our freedom as moral beings by being complicit in a system that systematically marginalizes brown people." Thus, Lebron's stated aim for this book is "racially relevant moral excellence" (xi).

Before proceeding through the chapters, it should be noted that although Lebron is avowedly motivated by his own and more broad experience of black