

Paul Katsafanas, *The Nietzschean Self: Moral Psychology, Agency, and the Unconscious*, (Oxford: Oxford University Press, 2016), xii + 292 pages. ISBN: 978-0-19-873710-0 (hbk.). Hardback: \$74.00

It is remarkable, given Nietzsche's insistence on the importance of psychology to philosophy, that engagement with his reflections on the mind has proven a predominantly recent affair. At the forefront of many such discussions has been Paul Katsafanas, whose latest book, *The Nietzschean Self (NS)*, offers an ambitious reconstruction of Nietzsche's moral psychology. According to Katsafanas, Nietzsche belongs to a philosophical tradition – represented by the likes of Plato, Aristotle, and the British sentimentalists – which approaches moral theory by first developing an “account of human nature” (p. 1). It is for this reason, we are told, that Nietzsche's works are replete with discussions of the role of reflection in action, the “genuine action/mere behaviour” distinction, the “structure of human motivation,” freedom, and responsibility (pp. 5-6), as well as the justification of ethical claims (p. 1).

Over the course of an impressively wide-ranging book, Katsafanas presents a systematic examination of the account of human nature we find in Nietzsche's writings. Structurally, *NS* involves two distinct movements. In the opening half, Katsafanas introduces the central components of Nietzsche's psychological terminology, and ascribes to him several substantial theoretical claims. This begins in earnest with a discussion of the unconscious (ch. 2). Katsafanas claims that, for Nietzsche, conscious and unconscious states are different in *kind*, the former having conceptual content, and the latter nonconceptual content (p. 46). The putative alignment of the conscious/unconscious and conceptual/nonconceptual distinctions is then employed to explain why Nietzsche takes conscious states to be *falsifying* (ch. 3). Consciousness, Katsafanas writes, “falsifies precisely because it is conceptual” (p. 72): In order to become conscious, the content of an unconscious state has to be interpreted – “pressed into a conceptual form.” This, it is claimed, involves both a simplification and generalization of unconscious content, which facilitate various types of falsification (pp. 55-62).

The next two chapters introduce the most distinctive feature of Nietzsche's psychological tool-kit: the *drive*. Making sense of Nietzsche's remarks on drives is a challenge. He writes of them interpreting and evaluating, and accounts for a number of complex phenomena in terms of their interaction. In an illuminating discussion (ch. 4), Katsafanas argues that Nietzsche follows Schopenhauer in conceiving of drives as “dispositions that

induce affective orientations” (p. 94). Drives interpret and evaluate the environment as they elicit behaviour by “coloring our view of the world, by generating perceptual saliences, by influencing our emotions and other attitudes, by fostering desires” (p. 97). Katsafanas develops this further (ch. 5), providing a detailed discussion of the relationship between drives and conscious evaluative judgments. As well as generating perceptual saliences, etc., drives, we are told, also “generate thoughts about justification” (p. 126). The resulting account suggests that an individual values X if they have a “drive-induced affective orientation toward X,” and do not disapprove of this orientation (p. 120).

This brings us to the second half of *NS*. Here, Katsafanas moves to Nietzsche’s treatments of various moral-psychological topics, in light of the framework established in the earlier parts of the book. In chapters 6 and 7, he addresses Nietzsche’s conception of willing, and the distinction between “*genuine action* and *mere behaviour*” (p. 165) respectively. Katsafanas argues that Nietzsche offers a subtle critique of the Kantian model of the will; one that, *contra* Kant, involves the rejection of “suspension” – the claim that when “an agent reflects on her motives for A-ing, she suspends the influence of the motives upon which she is reflecting” (p. 144). Nietzsche’s considered view on willing, Katsafanas insists, is a “*vector* model,” on which the will is “simply one source of motivation among many others,” including drives, affects, and conscious decisions (pp. 160-1). Katsafanas then turns to the problem of distinguishing action from mere behaviour (ch. 7), which he claims Nietzsche addresses by reference to the notion of *unity*. Whereas extant approaches have stressed the idea of unity among drives, however, Katsafanas suggests that the unity Nietzsche has in mind is a *harmony* between one’s drives and one’s conscious thought (p. 184-5). Notable in these chapters is the extent to which the picture of agency for which Katsafanas argues affords a far greater role to consciousness than is often ascribed to Nietzsche.

Chapters 8 and 9 broach the questions of selfhood and freedom. Katsafanas follows a number of commentators in claiming that selfhood, for Nietzsche, is “honorific or aspirational” (p. 200). In contrast to the dominant view, on which Nietzschean selfhood is a matter of the robustness and integration of one’s structure of drives, Katsafanas links selfhood to “some form of independence in valuing” (p. 220). Distinguishing Nietzsche’s position from what he calls the “Romantic view,” on which selfhood is linked to the purging of the effects of enculturation (ch. 8), and a “Kantian view,” on which it is tied to autonomy (ch. 9), Katsafanas advances the claim that an individual “qualifies as a self to the extent that she revalues her values” (p. 209).

The concluding chapter of *NS* presents a (favourable) comparison of Nietzschean moral psychology with competing views derived from the works of Kant, Hume, and Aristotle. Nietzsche's writings offer a wealth of insight to rival the great historical moral psychologists, and Katsafanas should be applauded for his efforts to demonstrate what Nietzschean moral psychology has to offer. *NS* is often insightful, and will rightfully influence subsequent work on Nietzsche's account of the mind. There are, nevertheless, elements of the book about which I have significant reservations. The first concerns Katsafanas' claim that Nietzsche aligns consciousness and conceptualization. This is employed at subsequent points throughout *NS*, yet the argument provided in support of it seems questionable. Katsafanas suggests (correctly) that conscious thinking, for Nietzsche, will be conceptually-articulated. To get beyond this uncontroversial claim to the claim that *only* conscious states have conceptual content, however, requires more. Accordingly, a lot of heavy-lifting is done by the insistence that there can be no concepts "without words," which Katsafanas derives from a single *Nachlaß* passage (p. 25).

This claim, however, is ambiguous. It might be taken to convey that language is the only appropriate vehicle for conceptual content. Alternatively, one might construe the claim as being that "words" (i.e., linguistic abilities) are necessary for concept possession. The former would entail that unconscious states lacked conceptual content, as Katsafanas claims, since Nietzsche is clear that only conscious thinking "*takes place in words*" (F. Nietzsche, *The Gay Science*, (Cambridge: Cambridge University Press, 2001), §354). Yet it would prevent conscious perception from having conceptual content, which Katsafanas argues it does on Nietzsche's view (pp. 30-37). The latter construal, while compatible with conscious perceptual states having conceptual content, only gets as far as demonstrating that word-possessing creatures *alone* have conceptual mental content, not that only the conscious states of such creatures do. Given the importance of this claim to Katsafanas' interpretation of Nietzsche's view, this shortcoming has implications for the tenability of a number of further interpretations Katsafanas offers.

My second concern relates to the textual support for a number of Katsafanas' claims. On occasion, he moves very quickly, ascribing substantial views to Nietzsche on the basis of slim evidence. For instance, Katsafanas attributes to Nietzsche a demanding standard for concept possession, involving, *inter alia*, the ability to appreciate a concept's place in a broader scheme (p. 25; 35). Yet the evidence for this is effectively Nietzsche's claim that "individual philosophical concepts are not arbitrary, do not grow up on their own, but in connection and relationship with each other" (Nietzsche, F., *Beyond Good and Evil*, (Cambridge: Cambridge

University Press, 2001) §20). Why this should be taken to generate a condition on concept *possession*, however, is left unclear.

The cumulative effect of such instances is a picture perhaps more ambitious than it is robust. Nietzsche provides some extremely rich and detailed analyses of particular phenomena (such as bad conscience and guilt), and I agree with Katsafanas that such views fit together consistently (p. 7). Yet it remains the case that attempts to draw out a systematic account of human nature from Nietzsche's writings will run into gaps that a more systematically-inclined philosopher might have felt compelled to fill. At times in *NS*, it feels as if Katsafanas drifts from a reconstruction of Nietzsche's view into something more akin to the construction of a Nietzsche-inspired view, in order to address such gaps. As an example, Katsafanas sometimes utilizes hypothetical cases – like that of addiction (cf. pp. 113-114; 177-179) – to argue against certain readings of Nietzsche. Putting aside whether or not these constitute effective criticisms of the views considered, it is unclear why our intuitions about such cases should have *any* bearing on the view we ascribe to Nietzsche.

Despite the above concerns, there is much to praise in *NS*. The book is elegantly-written, with complex exegetical work, philosophical discussion, and historical scholarship woven neatly into an admirably clear and readable whole. Whilst I remain unconvinced that Katsafanas' reading of Nietzsche is entirely successful, *NS* makes a good case for the value of both Nietzsche scholars and moral psychologists engaging seriously with Nietzsche's writings on the mind.

Christopher Fowles  
University of Oxford  
christopher.fowles@philosophy.ox.ac.uk