



***The Nietzschean Self: Moral Psychology, Agency, and the Unconscious* by Paul Katsafanas (Review)**

Richard J. Elliott

The recent publication of Paul Katsafanas's *The Nietzschean Self* establishes a benchmark for conducting systematic studies of Nietzsche. By treating Nietzsche's sporadic and often *prima facie* contradictory remarks on human agency in this well-considered study, Katsafanas successfully posits Nietzsche as an important contributor to philosophical issues surrounding the questions of human agency.

Katsafanas claims that accounting for human nature is the prerequisite to understanding the human good. Instead of morality leading psychology by the nose and dictating its limits, Nietzsche's claim is that psychology should be treated as "queen of the sciences" (BGE 23). In this respect, the book wishes to make a Nietzschean contribution to the field of moral psychology by explicating and assessing Nietzsche's account of human nature (4). Katsafanas characterizes the book's argumentative structure into six interrelated topics: the reflective or unreflective character of action, the difference between action and mere behavior, values and evaluative judgements, the dynamics or structure of human motivation, Nietzsche's account of freedom, and questions of responsibility and social determination.

Although he acknowledges that Nietzsche's claims on a great number of topics appear as either undercooked or contradictory, Katsafanas believes that one can excavate a comprehensive account of the human self from amidst the aphorisms. This account deals with drives, motivational states, the perplexities of the connection between the conscious and the unconscious, claims regarding the self's bearing towards history and society, and more. While Nietzsche's rejects the cobweb-spinning of metaphysical systematizers, he does not reject sustained inquiry, Katsafanas argues (7).

A considerable merit of Katsafanas's account is how well it sets a distinctly Nietzschean moral psychology against the dominant Kantian, Humean, and Aristotelian accounts, not least because the Nietzschean account is the only one which can satisfactorily deal with the nature of unconscious mental states and their function. Katsafanas also boldly claims that the Nietzschean account of the self he outlines is "consonant with our best empirical and philosophical views" (9) and offers a more realistic picture of human psychology, one less constrained by the moralistic enterprise that underpins the dominant attempts at constructing models of agency.

Another merit of the book is its rich dialogue with rival contemporary interpretations of the themes which Katsafanas highlights within Nietzsche's work. Nearly every chapter gives exegetical space and critical commentary to other interpretations, setting them out clearly in their positions and comparing them to Katsafanas's own positions. Katsafanas helpfully sets aside dozens of pages for both summary and critical remarks on each aspect of Nietzsche's thought.

In chapters two through four, Katsafanas treats the nature of the unconscious, its bearing on conscious mental states, causal efficacy, and the motivational power of drives for Nietzsche. Katsafanas attempts to square up Nietzsche's claims that consciousness is both superficial and falsifying (14) with a coherent, distinctly Nietzschean model of agency. With the aim of identifying what exactly Nietzsche means by an unconscious mental state and how it relates or bears upon consciousness, Katsafanas leads us through a terse but interesting history of how unconscious mental states have been conceived, as well as the manner in which they have been understood as either underlying conscious states, or as in competition with conscious states.

With nods to the historical contributions on this matter from Leibniz, Fechner, von Helmholtz, Lange, Afrikan Spir, Freud, Schelling, Hartmann, Herder, and Schopenhauer, Katsafanas shows how Nietzsche's rejection of the causal efficacy of the 'Ego' has been mistakenly understood as a rejection of the causal efficacy of consciousness, understood as the sum totality of conscious mental states (22-23). Although Nietzsche casts aside a central Cartesian tenet of human agency, he does not discount the role of conscious thinking completely. Rather, Katsafanas argues, consciousness is essential for the communicability between humans. In this sense, consciousness is inherently tied to language, which Nietzsche identifies as having a falsifying influence.

When Katsafanas claims that the difference between conscious mental states and their unconscious counterparts lies in the former possessing conceptual content and the latter possessing non-conceptual content, we see how Nietzsche ties this in with his claim at BGE 268 that words function as “acoustical signs for concepts.” When Schopenhauer writes that language allows for “the summarizing into one concept of what is common” (WWR I, 37), we see a distinct echo of this in Nietzsche, who claims that this capacity is one which limits or dulls the richness of mental life, something which previous accounts of moral psychology have not acknowledged. Yet as Katsafanas persuasively argues, Nietzsche distinguishes himself from Schopenhauer by offering two types of perception: perceptions with conceptual content, and perceptions with non-conceptual content (31). Katsafanas’s excursus on the differences between conceptual content and non-conceptual content in perceptions, and justification for aligning the conscious/unconscious distinction with the conceptual/non-conceptual distinction (30 – 46), are well worth the labour of reading, not least for his impressive and largely convincing responses to contemporary secondary literature on these topics. Unlike many contemporary philosophical discussions about the unconscious which set a rigid divide between conscious and unconscious states, Katsafanas argues that Nietzsche sees their relation as “more continuous” (46), on a spectrum of degree of awareness.

Katsafanas ties his claim that conceptual content structures conscious experience to Nietzsche’s much-debated perspectivism (52–54). Although Nietzsche accepts a qualified version of the Kantian conceptual scheme, Katsafanas argues that he distinguishes himself from Kant with his notion that the conceptual structuring of conscious experience is social or historically fluid. According to Katsafanas, this is what Nietzsche’s perspectivism amounts to (53). The perspective of the human effects the transmission of unconscious mental states into conscious ones. Here, Katsafanas gives the illuminating example of bad conscience from the second essay of Nietzsche’s *Genealogy*, wherein the unconscious state of bad conscience is manifest in the form of conscious states as guilt (57-63). And while Katsafanas does not outline a positive claim regarding the structural relations of the unconscious “because any expression of these relations would press them into a conceptual structure, thereby falsifying them” (65), he does nod in the direction of Freud in arguing that like him, Nietzsche thinks that the regulation of unconscious mental states must be different from conscious ones. While unconscious processes represent causal relations in themselves, conscious processes represent only

“familiar” or commonly “intelligible” causal relations (67). Yet Nietzsche’s understanding that unconscious desires cannot be articulated by linguistic means distinguishes him from Freud (70). This seems right to me, so much the worse for psychoanalytic appropriations of Nietzsche.

With regards to Nietzsche’s drive psychology, Katsafanas distances himself from reductionist accounts of drives as mere physiological states, but likewise rejects interpretations that would remain open to homunculus fallacies. Katsafanas qualifies that Nietzsche’s agential language about the nature of drives shouldn’t be taken literally so as to identify drives themselves as agentially conscious, and normatively sensitive (78-82), but rather should be considered to be agentially significant as “embodied drives” (97-98). While also detaching himself from identifying drives as “mere” urges or dispositions (84), Katsafanas offers a qualified reading that understands Nietzsche’s drives as dispositions that “induce affective [evaluative] orientations in the agent” (86).

By drawing on Schopenhauer’s discussion of the reproductive drive’s capacity to produce desires and so to influence the agent’s response to phenomena and other agents, Katsafanas argues that we should make sense of the drive-affect discourse at work in the agent in such a way for Nietzsche’s account of all drives: Drives make particular phenomena salient and others peripheral, which shapes both our attitudes and even influences the content of experience itself (94). There is much to find agreeable in Katsafanas’s treatment of these topics. There is also much to gain from his exposition of the dynamics of drives, especially the characteristic distinction he draws between a drive’s aim and its object (101-102).

The fifth chapter builds upon this analysis, in the service of understanding why drives possess evaluative significance for Nietzsche. According to Katsafanas, drives “generate thoughts about justification,” thus providing strong inclinations to consider the drive’s end as valuable for the agent (108). Crucially for Katsafanas, Nietzsche does not equate values with the aims of drives in themselves, since drives can be merely “cravings” (111), in the manner that the religious ascetic still possesses a sex drive, but doesn’t value sexual activity more than as a mere “responsive disposition” (113). Once again, Katsafanas’s engagement with the contemporary literature is commendable in this chapter, particularly in relation to Nietzsche’s understanding of the affects and the requisite

justification the agent must bestow on them for it to be constituted as a “value,” properly understood (120).

The sixth chapter deals specifically with Nietzsche’s account of agency. For Nietzsche, motives are not causally indeterminate as in the traditional (Kantian) account. Weighing in on the reasons versus causes debate on behalf of Nietzsche, Katsafanas claims, I think correctly, that reasons are incapable of quelling motives as possessing causal salience in dispositions to act (136). At the same time, Katsafanas does not rule out conscious deliberation as possessing a causally efficacious role as well (148), citing Nietzsche’s claim that conscious thought can alter or redirect the motivational dispositions of the affects. As well as offering a detailed account of the development of Nietzsche’s thoughts on willing, Katsafanas offers the helpful comparison with Kant’s account of agency. Katsafanas posits what he calls a “vector” account of willing as true to Nietzsche’s, one which doesn’t generate *causa sui*-type forces, but rather modifies forces or drives that are perpetually at work (160).

It is the power to consciously make decisions that marks the distinction between “strong” and “weak” wills, Katsafanas argues. In the seventh chapter, Katsafanas discusses this in terms of Nietzsche’s conception of unity and responsibility, by addressing the difference for Nietzsche between actions and mere behavior (164-165). Eschewing the common equation of freedom with unity in Nietzsche, Katsafanas argues that unity is best understood as a particular kind of harmonious relation between one’s drives and one’s conscious thought (193): Katsafanas argues that an agent counts as unified in Nietzsche’s sense if “further knowledge of the drives and affects that figure in [his] etiology would not undermine this affirmation of [his particular constitutive relation of drives and affects]” (192).

Building upon the arguments of the antecedent chapters, chapters 8 and 9 make several arguments regarding Nietzsche’s understanding of selfhood, the relation between society and the agent, and freedom. Katsafanas argues that we should understand the term “self” “to refer to those who bear an appropriate relation to their culture” (198). Katsafanas describes the great individual as one who “not only embodies this new ideal, but also plays a transformative role, shifting groups of entire societies toward new hierarchies of value” (ibid.).

Katsafanas also critiques John Richardson's claim that "social selection instils behavioral dispositions that are contrary to an individual's own interest" (211); according to Katsafanas's Nietzsche, there are no pre-social drives, and human nature is malleable. Katsafanas argues that Nietzsche does not endorse the Romantic claim that genuine selfhood is achieved by freeing ourselves from the dominant social norms and values of one's time, place, and culture. But this is questionable. The textual evidence that Katsafanas employs here in support of his position is some of the most stretched in the book, especially given Nietzsche's discussions elsewhere of the differences between "lambs" and "birds of prey" (GM I: 13) and his explicit endorsements of amorality in the higher, exemplary types of individual. What if the critical assessments of one's values, that Katsafanas identifies as being the qualificatory standard of genuine selfhood, meant abandoning the dominant social moral norms of the day? Following Katsafanas's argumentation in chapter 8, this question doesn't seem to receive an adequate response. Katsafanas argues that Nietzsche adopts his own brand of Hegelian *Sittlichkeit*, but this interpretation does not leave room for Nietzsche's imperative to initiate a truly radical critique of dominant values amongst exemplary individuals or the prospect of their rejection that must remain a possibility for the option for true flourishing.

Further, Katsafanas draws upon Nietzsche's discussions of the conflict between social customs and the drives, but if no drives exist pre-socially, how can such a conflict arise? How could Nietzsche conceivably describe Judeo-Christian societies as possessing an "unnatural" morality that "runs counter to sense, instinct, nature, animal" (GM II: 24) if the true nature of drives was constructed by their social circumstance? Katsafanas attempts to answer this problem by saying that "the fact that there are many conflicts between particular customs and particular drives does not entail that we should accept a custom/drive dichotomy, or that we should see the drives as things that can be understood as having pre-social aims" (214). But even the fact that conflicts are possible between customs and drives even in particular contexts demonstrates that there exists a categorical distinction between the two. There must be a "natural" state of the drives that allows for them to come into conflict once posited at the level of the social. To employ Katsafanas's own aim/object distinction from Chapter Four, the constitutive aspects of human drives bear a relation to societal customs, rather than being the same thing as them.

These claims about the drives also appear to sit uneasily with Katsafanas's claim in chapter 9 that the free individual can set goals liberated from morality, and that "he can regulate his behavior without reliance on external factors" (229). Katsafanas understands the higher individual's drives as intrinsically socially bound, but he also describes the "sovereign individual" as being detached from external influences (229-231). Katsafanas attacks so-called "radical subjectivist" readings of Nietzsche for being "empty" (233-4), but his own position seems inescapably committed to a quasi-Hegelian view of Nietzsche, according to which exemplars accept their relations to institutions and social practices. Even as he emphasizes how "radical" Nietzsche's critique of moral and social norms is in comparison to Hegel's, one must ask: how radical can Katsafanas's Nietzsche really be? The ninth chapter ends with a notion of Nietzschean "immoralism" which commits the immoralist to the contradictory task of "reassess[ing] even [his] most basic values" (253) even as he recognizes that his drives could not be otherwise than how they are manifested in the realm of the social.

The final sentence in Katsafanas's book outlines his hope for what his book might have accomplished: that "there is a philosophically fruitful alternative" to dominant philosophical discussions about agency in the form of a Nietzschean model (279). In this respect, I think Katsafanas is successful: His contribution here to both Nietzsche studies and to contemporary moral psychology is original, fruitful and innovative. As should be expected of a book that so thoroughly gives treatment to many of the core topics and questions arising from within Nietzsche's work, there are questions to be raised as to the accuracy of certain aspects of Katsafanas's reading of Nietzsche, but the originality in the treatment of the topics covered in Katsafanas's book is a great service to Nietzsche studies.

Richard J. Elliott

Fall 2016