

Maudemarie Clark and David Dudrick, *The Soul of Nietzsche's "Beyond Good and Evil"*.
Cambridge: Cambridge University Press, 2012.

Paul Katsafanas

Draft of May 27, 2013

Forthcoming in *Journal of Nietzsche Studies*

If Clark and Dudrick have their way, gone will be the days of breezy writings on Nietzsche that recruit a phrase from here, a paragraph from there, and construct an interpretation from the resultant *mélange*. Clark and Dudrick advocate a meticulous, line-by-line study of Nietzsche's texts, with painstaking attention not only to the broader context of Nietzsche's claims, but even to the precise intent of the images and metaphors that he employs. Here, we find a level of textual scrutiny and careful consideration of context that has been largely absent in Nietzsche scholarship. To get a flavor of the book, consider the fact that Clark and Dudrick spend no less than 63 pages on the preface and first four sections of BGE. Indeed, there's a sense in which the first half of the book is devoted to an analysis of one metaphor. The detail and precision is admirable.

The book is not only meticulous—it is exceptionally rich. Though focused on the first part of BGE, it discusses a host of topics ranging from naturalism, physiology, philosophical psychology, will to power, realism and anti-realism, positivism, sensualism, Kant and Hume on causality, Spir on dogmatism and the relations between causal and normative claims, and Plato on the soul. Many of these discussions are illuminating, introducing intriguing new takes on these debates.

The book falls into two parts. The first part focuses on a metaphor that Nietzsche employs in the Preface: there, Nietzsche remarks on a “magnificent tension of the spirit” that “characterize[s] the current situation of philosophy” (11). Clark and Dudrick argue that we should interpret this tension as generated by two competing forces, which they call the “will to truth” and the “will to value” (12). They argue against readings of Nietzsche that strive to eliminate or mitigate one of these wills, instead proposing that Nietzsche sees a way in which these wills can coexist productively (113-39). I’ll spend Section One of this review investigating these claims. In particular, I’m skeptical that Nietzsche would recognize a distinct, independent will to value, and accordingly I offer a different reading of the relevant passages.

The second part of the book analyzes Nietzsche’s philosophical psychology, with a focus on his discussions of willing, drives, will to power, and the structure of the self (or “soul,” as Clark and Dudrick prefer). There, they argue that Nietzsche’s will to power theory is “a doctrine of what constitutes the human soul, what makes us persons or selves, hence what differentiates humans from other animals” (139). In particular, they maintain that the Nietzschean self comprises various drives that stand in “normative relationships” to one another. As they put it, “the human soul is not a naturalistic entity, but neither it is a metaphysical one. It is a normative entity, which exists only in and through the space of reasons” (139). Sections Two and Three of this review will address these claims. In Section Two, I argue that Clark and Dudrick’s analysis of Nietzschean drives is philosophically and textually problematic. Finally, Section Three investigates their claim that Nietzsche understands the self as a “normative ordering” of drives, which they distinguish from a “causal ordering.” I raise some questions about the cogency of this distinction.

1. The will to value versus the will to truth

As I mentioned above, the preface of BGE discusses a “magnificent tension of the spirit,” and Clark and Dudrick take BGE I to be an explication of this tension. They suggest that “the future Nietzsche envisions for philosophy... depends on the proper resolution of this tension” (11). If this is right, then understanding the tension will be crucial. The “key to understanding this metaphor,” they write, “is to recognize that the tension in question must be produced by a conflict between two different forces... We find them in the will to truth and what we call ‘the will to value’” (12). Accordingly, an overriding idea in Clark and Dudrick’s book is that Nietzsche is concerned with a tension between a “will to value” and a “will to truth.”

I have two concerns about this claim. First, the contention that the will to value is in conflict with the will to truth seems insufficiently precise. Second, it’s not clear why Clark and Dudrick take Nietzsche to countenance a discrete will to value.

Start with the first problem. Given the way in which Clark and Dudrick define the will to value and will to truth, it’s unclear how they can be in conflict with one another. Clark and Dudrick write that “the will to truth aims at believing only what corresponds to the way the world actually is,” whereas “the will to value aims to represent the world in terms of what is valuable, in terms of what it would be good for the world to be” (37).¹ However, Clark and Dudrick note that the will to truth, so defined, constitutes an evaluation: they write that Nietzsche “understands the *will* to truth as a *commitment* to the value of truth” (35). But given these definitions, doesn’t the will to truth turn out to be a *species* of will to

¹ Though see p. 44, where Clark and Dudrick define will to value and will to truth differently: there, we are told that the will to truth is “the will to see that world as it *is*” and the will to value is “the will to see the world in a way that accords with [one’s] values.”

value? An agent has a will to truth if she is committed to the value of truth; so, an agent has a will to truth if her “will to value” is directed in a certain way—at truth. This entails that the “magnificent tension” between will to truth and will to value would actually be a tension *within* the will to value. It would be a tension between *particular values*, rather than between value and something else.

This brings us to the second concern. What exactly is the will to value supposed to be, and where in Nietzsche’s texts do Clark and Dudrick find evidence for it? As far as I can tell, Clark and Dudrick rest their case for a “will to value” on BGE 2. They note that BGE 2 investigates the motives that drive philosophers’ metaphysical theorizing, and claim “it is here that [Nietzsche] introduces the will to value” (37). In particular, they read BGE 2 as claiming that “a ‘valuation’ stands behind one particular species of philosophy, metaphysical philosophy,” and they see the following sections as extending this point: “BGE 5-8 argue that some such valuation stands behind all philosophy” (42). Thus, Clark and Dudrick move from the idea that

(1) whenever a person engages in philosophy, some value motivates this pursuit
to the claim that

(2) the best explanation for (1) is that there exists a will to value.

The idea, I suppose, is that the will to value finds expression in philosophizing (as well as other activities), and therefore explains (1).

My objection is that there are *many* ways to explain (1) without appealing to (2). For example: perhaps what motivates Kant’s philosophizing is a valuation of religion. And perhaps what drives Schopenhauer’s is a disvaluation of suffering. And perhaps what drives Feuerbach’s is a valuation of scientific progress. And so on. In other words, perhaps each

philosopher is driven by an antecedent commitment to some *particular value*, rather than a commitment to an amorphous “will to value.”

It might seem that there is no great difference between claiming (a) various agents have various, particular values, and (b) agents have a will to value. However, this is a mistake: (a) and (b) are importantly distinct. On reflection, the idea of an independent “will to value” is very odd. It suggests that human beings aim at valuing things, and go about looking for ways to fulfill this aim. That is, it suggests that I want to stand in relations of reverence and respect to something, so I go about trying to fill in the blank. But an alternative idea is that I don’t have a blank aim of valuing things. Rather, I value *particulars*, and I value them *because I have a drive toward them*. I have a drive toward sexual activity, and so value sex; I have a drive toward knowledge, and so value knowledge; I have a drive toward sociality, and hence value social activity. This view, it seems to me, fits much better with Nietzsche’s claim that each drive is a valuation. Just as the will to truth involves a valuation of truth, so too a will to sexuality involves a valuation of sexuality, a will to aggression involves a valuation of aggression, and so on. But none of this implies that there is some independent, discrete will to value that then finds expression in various particular values.

Why does this matter? What’s wrong with moving from the idea that I value X, Y, and Z to the idea that I have a drive toward valuing? Well, it’s a bit like concluding, from the fact that I am committed to taking the means toward knowledge, sex, food, sociality, that I have a blank aim of taking means to ends, and go about adopting ends so as to fulfill this aim. That’s backwards. Rather, I’m attracted to particular things, and am thus inclined to take the means toward them. Just so, I suggest, with the will to value. The talk of a single will to value suggests that I have a blank aim of valuing things and go about looking for ways to fulfill it. But this seems phenomenologically implausible and philosophically problematic.

If we had strong textual support for the idea that Nietzsche countenances a will to value, then we might set these problems aside; but, as I've indicated above, the textual evidence is exceedingly thin.

Suppose this is right; suppose there is no one will to value, just many values. Then the “magnificent tension” looks very different. It's a tension *within* the will to truth, which tries to deny its own evaluative status. This, after all, is the tension that Nietzsche explores in the *Genealogy*: the valuation of truth, when taken seriously enough, leads us to question our grounds for valuing truth. No “will to value” is involved in this story, just one particular value.

2. Drives as homunculi

The second half of Clark and Dudrick's book is concerned with philosophical psychology, and I'll focus the remainder of the review on this material. I want to start with their analysis of Nietzsche's notion of drive [*Trieb* or *Instinkt*]. When explaining a person's actions, values, conscious thinking, and indeed the structure of the agent's self, Nietzsche appeals to drives. But when we ask what, exactly, a drive is, puzzles arise. Nietzsche tells us that drives “adopt perspectives,” “interpret the world,” and “evaluate.”² For example, in BGE 6, Nietzsche writes that every drive “would be only too glad to present just *itself* as the ultimate goal of existence and as the legitimate *master* of all the other drives. For every *drive* is desirous of ruling; and it is as *such* that it tries to philosophize.” This language of valuing, adopting perspectives, presenting oneself as master, and philosophizing is ordinarily used only with full-fledged agents. For this reason, some commentators have interpreted

² For some examples, see KSA 12:1[58], WP 481, 260, 567. xxx

Nietzschean drives as homunculi—sub-personal entities that have the properties of agents (see Poellner 1995). Clark and Dudrick wholeheartedly embrace this interpretation, arguing at length that Nietzschean drives are homunculi (see especially 196-8).

To see this, consider the fact that Clark and Dudrick attribute extremely sophisticated capacities to drives. They tell us that drives are aware of one another and try to “to prevent other drives from getting what they want” (146). They claim that each drive “systematically develops and defends *an account* of reality from its point of view” (146). They imagine drives “commanding other drives to carry out certain actions” (183) and “presenting themselves to the other drives as having political authority, as having the authority to speak for the whole ‘commonwealth’... In taking this stance toward the other drives, they ‘experience’ themselves as superior to them” (183). They suggest that just as a political ruler in a state rules not just because she is strongest, but because she is “recognized as having the authority to rule,” so too “one drive has a higher rank than another not in virtue of causal efficaciousness...but in virtue of being recognized as having a *right* to win” (150). In short: drives are homunculi. They are miniature agents who communicate with one another, develop political orderings, perceive authority relations, have plans and strategies, have experiences, develop accounts, and so on.

I’ve argued elsewhere that homuncular readings of Nietzschean drives are indefensible; instead, we should understand drives as dispositions to token affective or evaluative orientations and patterns of activity.³ When Nietzsche claims that drives “philosophize”, evaluate, interpret, and so on, he is remarking on the way in which an embodied drive generates affective or evaluative orientations in an agent; he is not committed to the idea that the drive *considered in isolation from the whole agent* has these

³ Paul Katsafanas, “Nietzsche’s Philosophical Psychology,” in Ken Gemes and John Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford University Press, 2013), 727-55.

experiences. Here, I want to reiterate some of these concerns. I think there are compelling philosophical and textual reasons for rejecting the homuncular reading of drives.

Let's start with the philosophical problems. First, Clark and Dudrick attribute experiences, perceptions, the capacity to issue and obey commands, and the capacity to recognize political authority to drives. It's hard to see how this is possible without the drives' being self-conscious. But surely Nietzsche cannot imagine that each human being comprises a multitude of self-conscious entities.

Second, it's not obvious how the homuncular view of drives could have any explanatory power. For example, consider Nietzsche's efforts to explain conscious agency in terms of drives. If drives are themselves conscious agents, what exactly is being explained here? We want an explanation of conscious agency, and we are told to understand a person's conscious agency as a manifestation of the conscious agency of various drives. This is hardly informative. Rather than explaining agency and selfhood, it simply shifts the problematic terms about, from the level of persons to the level of drives.

Clark and Dudrick attempt to respond to these concerns in two ways. First, they claim that homunculi are problematic only if they are just as complex as the whole organism. They write that "because these activities are simpler than the one for which they are supposed to account, there is no problematic circularity here" (198). The activity for which drives are supposed to account is "the agency characteristic of a human being" (197). So, Clark and Dudrick claim, commanding, obeying, and recognizing political authority are simpler than the agency characteristic of us.

However, this claim is difficult to believe. In what sense are commanding, obeying, and recognizing authority "simpler" than other manifestations of human agency? Commanding and obeying, for example, require very robust cognitive processes. After all,

they require, at minimum, the presence of consciousness: one entity must perceive another as commanding, understand the nature of this command, be motivated to bend itself to this agent's command, and then carry out the command. How could a subpersonal psychological entity, a drive, do this?

This brings us to the second aspect of Clark and Dudrick's response. They claim that commanding and obeying cannot be understood in merely causal terms: as they put it, we must distinguish commanding and obeying from "mere physiological strength of the drives" and "brute causal strength" (198). Thus, when drive A commands drive B, this involves more than A simply overpowering B. Instead, drives issue commands when they exert "*political* authority" (198). Drive A commands drive B when drive B recognizes and responds appropriately to drive A's authority.

What is the difference between causal strength and political authority? That is, how do we distinguish between drive B's being overpowered by A and B's submitting to A's authority? In order to elucidate this distinction, and to respond to the concern that drives are being treated as self-conscious agents, Clark and Dudrick point out that contemporary scientists recognize "political" relations or "dominance hierarchies" among certain animals, such as wolves and chimps (198). These relations, Clark and Dudrick claim, are not merely causal; they are normative. And their idea seems to be that if wolves and chimps can stand in these normative relations, so too can drives.

Suppose Clark and Dudrick are right that the relations among drives are analogous to the relations among wolves and chimps. Clark and Dudrick take this analogy to assuage the concern that drives are being treated as full-fledged agents. I would interpret it in exactly the opposite way: it makes it clear just how dire the problem is. Wolves and chimps stand at the heights of cognitive sophistication among animals. Insofar as we attribute wolves and

chimps, but not ants and bees, the capacity to recognize political relationships, we recognize that genuine commanding and obeying requires tremendously more than mere ordered action. Ants, for example, have complex, hierarchically ordered societies; but we do not suppose that ants actually recognize political authority. Certain scientists do, on the other hand, attribute that capacity to chimps. If subpersonal drives are supposed to have this degree of cognitive sophistication—if psychological processes inside of me are somehow supposed to have the mental capacities of a chimp—then Nietzsche’s theory belongs in the rubble heap of outlandish nineteenth-century biological speculations.

Clark and Dudrick do have a response. They claim that contemporary scientists take the social hierarchies exhibited by animals such as wolves and chimps not to require that “the animals in question *take themselves* to form a political order; their ‘conscious motives and intentions’ need not concern their political standing” (199). Instead, the claim that animals stand in political relations “means only that the behavior of an individual is best explained not only in terms of his brute strength relative to his fellows but in terms of his rank in the social order. In saying that drives form a political order, then, Nietzsche need not take them to be *conscious* of their political situation—he need not take them to be conscious *at all*” (199).

However, Clark and Dudrick’s final remark does not follow from their interpretation of the empirical data: the data that they cite claims that social hierarchies don’t require consciousness *of political order*, but from this it hardly follows that social hierarchies don’t require consciousness *at all*. A chimp that defers to an authority figure may not be conscious of standing in a political order, but it is clearly conscious *of something* (the menacing posture of the domineering chimp, for example). So the empirical data in no way supports Clark and Dudrick’s claim that drives could stand in political relations without being conscious.

In sum: Clark and Dudrick's analogy between drives and chimps does not help their case at all. Rather than assuaging the concern that drives are being attributed extremely sophisticated mental capacities, the analogy drives home the concern.

Of course, it's possible that Nietzsche fell into error and endorsed this rather dubious theory. However, this brings us to a decisive textual problem with attributing the homuncular view to Nietzsche: it is extremely difficult to reconcile this interpretation with Nietzsche's other commitments. Nietzsche makes it quite clear that he wants to rethink our notion of the self:

And as for the Ego! That has become a fable, a fiction, a play on words: it has altogether ceased to think, feel, or will! (TI vi.3)

To babble about "unity," "soul," "person," this we have forbidden: with such hypotheses one only complicates the problem. (KSA 11:37[4])

These passages question our ordinary understanding of the self. As Clark and Dudrick note, Nietzsche argues that once we recognize that the self harbors multiple drives, we must rethink the nature of the self. But if drives are homunculi, then Nietzsche's rethinking of the self is a rather modest affair: Nietzsche would simply be claiming that there are many more selves than we thought. In other words, the homuncular interpretation assumes that we already have a coherent concept of selfhood, and are simply mistaken as to which entities instantiate this concept: we thought that whole persons instantiated selfhood, but we find that parts of persons – drives – instantiate selfhood.

This interpretation is dubious. Nietzsche seems to be claiming, not simply that we have applied the concept of selfhood to the wrong entity (person rather than drive), but that

we do not even possess a coherent concept of selfhood. In other words, Nietzsche is not simply claiming that there are more selves than we think there are; instead, he is claiming that we have a mistaken conception of selfhood. He wants to transform our notion of selfhood, not simply to apply the notion in a more profligate fashion. I take it that this is part of what Nietzsche means when he writes that drives “are not soul-atoms” (KSA 11:37[4]/WLN 30).

Let me close by pointing out that Clark and Dudrick interpret the “atomistic need” which Nietzsche rejects in BGE 12 as “the need for a unit at the microlevel that is a smaller version of the things we are familiar with at the macrolevel” (161). Though they see that atomistic need as resting on “plebian assumption[s]” and grammatical errors (161), it’s hard to avoid noticing that Clark and Dudrick’s homuncular drives are paradigms of this need—for what are their drives but micro-versions of ordinary agents?

3. The alleged conflict between causal and normative accounts of the self

These reflections lead us to our final concern. Clark and Dudrick rely on the idea that the self comprises drives that are normatively or political ordered, rather than causally ordered (139). They reject the idea that “the causal order” of the drives explains the type of person one is, why the person holds her values, and why she holds her theoretical views (149). Instead, they suggest that Nietzsche “takes the soul to be ‘the political order of the drives and affects’... The political order is not just a causal order, but a normative one” (150). Although they acknowledge that Nietzsche “never comes right out and says that it [the ordering of drives] is ... a normative rather than a merely causal order” (155), they claim that we can make sense of his claims only by interpreting him in this way.

I confess to being unsure what, exactly, Clark and Dudrick mean by “normative order” and why they take the claim that some domain has a normative order to be incompatible with the claim that it has a causal order—why think the two kinds of “orders” are in conflict with one another, rather than just different vocabularies employed in describing different features of the same thing? I lack the space to pursue this concern in detail here, though, so let me raise more restricted points.

First, whatever we think of the tenability of the causal/normative distinction in general, it seems to me odd to interpret Nietzsche as endorsing any such distinction. Put a bit differently, it’s surprising to think that Nietzsche would endorse a strong fact/value distinction— isn’t he constantly emphasizing the ways in which allegedly non-evaluative, affectively neutral descriptions actually presuppose and contain evaluations? Where does one find Nietzsche countenancing such a distinction? To the extent that he writes about such a distinction, he seems to me extremely critical of it.

Clark and Dudrick make much of Nietzsche’s rejection of “mechanism”; they take his rejection of mechanism to involve an endorsement of the causal/normative distinction. However, I think it’s clear that Nietzsche thinks the mechanists go wrong not because they focus solely on the causal and ignore a separate normative realm, but because they don’t see that our account of the causal is already, and pervasively, normative. The mechanists’ conception of causes, substances, subjects, and so on already contain—or so Nietzsche wants to claim—evaluations.

Put a bit differently, Clark and Dudrick’s distinction between the causal and the normative seems to presuppose something like this:

- (A) We can offer a non-evaluative or non-normative description of the world, but if we do we will miss something crucial. The mechanists and other clumsy

naturalists go wrong because their descriptions of the world leave out values and reasons.

Whereas I suggest that Nietzsche's point is rather

(B) We cannot offer a non-evaluative or non-normative description of the world, because our descriptions ineluctably presuppose or constitute evaluations. The mechanists and other clumsy naturalists go wrong because they do not realize that their descriptions of the world, which purport to be purely factual, harbor or constitute evaluations.

But suppose I'm wrong; suppose Nietzsche actually does accept (A). This brings us to a second problem. Why exactly is Nietzsche supposed to deny that the structure of the agent's drives can be given a causal explanation? Clark and Dudrick write that "the drives are arranged not merely in a causal order but in a political one" (175). They claim that the political order cannot be given a causal explanation, because it is not merely "an order of strength" (189). That is, we cannot explain the political order among the drives merely by appealing to which drive is strongest. It seems, then, that Clark and Dudrick rely on the following argument:

1. The drives stand in various relationships.
2. These relationships between drives cannot be explained merely by characterizing certain drives as stronger than others.
3. Therefore, the relationships between drives cannot be explained in causal terms.
4. Therefore, the relationships between drives must be explained in normative terms.

If this is the argument, though, it is unconvincing, for claim (3) does not follow from claim (2). Suppose I want to give an analysis of a human political organization, such as a university. It's true that an analysis that merely says "the president is strongest, the dean a bit weaker, the faculty weaker still..." won't be very informative and won't capture the interactions between these constituents. But why is that supposed to be the only option? Causal explanations can be much more refined than this; they can cite a much broader range of factors than mere strength. The fact that a *crude* causal explanation fails does not entail that *all* causal explanations fail.

After all, Nietzsche makes it very clear that he doesn't believe facts about drives alone are going to be enough to explain a person's behavior. Just look at *Ecce Homo*: there, we're told that Nietzsche's diet, the climates he lived in, his childhood experiences, and so on exerted a decisive influence in making him who he was. So a full explanation of a person's behavior is going to include not just drives, but environmental and physiological factors, as well as experiences and memories.

Why not try something similar with drives? It's probably true that trying to explain the relationship between, say, the knowledge drive and the sex drive merely in terms of which is stronger won't be very informative. But that's not the only option. We could instead say that the knowledge drive is dominant in the sense that it redirects other drives, structures the person's deliberations as well as her conceptualization of her environment, tokens affective orientations in certain circumstances, leads to certain patterns of behavior, and so on. Even if the simple causal story fails, this more sophisticated causal story might succeed.

Although I've raised some objections to Clark and Dudrick's analyses of the will to value, drives, and the causal/normative distinction, I want to close by emphasizing just how valuable I take their book to be. It is an unusually insightful work, full of nuanced analyses, valuable reflections, and challenging objections to competing interpretations. Everyone with an interest in Nietzsche will need to read it.

Paul Katsafanas

Boston University

pkatsa@bu.edu