

Nietzsche on Agency and Self-Ignorance

Paul Katsafanas

Penultimate draft. Final draft will appear in *International Studies in Philosophy* volume 40.3.

Presented at the 2006 Central APA North American Nietzsche Society session.

Nietzsche labels the widespread belief that agents are cognizant of their own actions “the universal madness.” For “the opposite is precisely the naked reality demonstrated daily and hourly from time immemorial! ... Actions are *never* what they appear to be... all actions are essentially unknown” (D 116).¹ Indeed, Nietzsche seems to claim that we *never* know what we are doing: “we are necessarily strangers to ourselves, we do not comprehend ourselves, we *have* to misunderstand ourselves, for us the law ‘each is furthest from himself’ applies to all eternity” (GM P 1). He never tires of emphasizing that “everything about [an action] that can be seen, known, ‘conscious,’ still belongs to its surface and skin—which, like every skin, betrays something but conceals even more” (BGE 32).

Although it will take some exegetical work to understand Nietzsche’s claims, the suggestion that we are ignorant of our own actions is profoundly disturbing. For it is part of our ordinary conception of agency that agency requires self-knowledge. Many contemporary philosophers have argued that if an agent is ignorant of his action, then the appearance of agency is illusory.² So, as Michael Bratman puts it,

An agent moved by desires of which he is unaware, or on which he is incapable of reflecting, or from whose role in action he is, as we sometimes say, estranged, seems himself less the source of the activity than a locus of forces. (Bratman 2001, 311-12)

Accordingly, Bratman suggests that we should distinguish between these degenerate cases of “merely motivated behavior” and cases of “full-blown” or “genuine” agency (Bratman 2001, 312).

Nietzsche’s emphasis on self-ignorance threatens to undermine the possibility of genuine agency. Does Nietzsche intend us to conclude that we are never truly in control of our actions, that we are mere loci of forces, that our conscious thoughts and reflective decisions are like so much flotsam buffeted about by unconscious forces?

I. NIETZSCHE’S ARGUMENTS

To answer these questions, we must begin by examining Nietzsche’s arguments. The general idea that self-ignorance undermines agency must be made more precise, in two respects. First, it is not clear how, exactly, agency requires self-knowledge. Second, it is not clear just how extensive Nietzsche believes this self-ignorance is.

To begin with the first point, agency is typically linked to self-knowledge through a third term: *control*. One of the few uncontroversial things we can say about agency is this: if a person is not in control of her behavior, then her claim to being an *agent* is undermined. Moreover, an agent’s being in control of her behavior seems to require, as a minimal condition, that the agent have some conception of what she is doing.

Nietzsche alludes to the connection between agency, self-knowledge, and control in a somewhat obscure passage:

'I have no idea what I am *doing*! I have no idea what I *ought to do*!' – You are right, but be sure of this: *you are being done!* [du wirst gethan!] at every moment! Mankind has in all ages confused the active and the passive: it is their everlasting grammatical blunder. (D 120)

Although this passage appears very natural, it contains a substantial philosophical thesis. It moves from a claim about *self-knowledge* (I have no idea what I am doing) to a claim about passivity (you are not *doing*, you are *being done*). It therefore implicitly commits Nietzsche to the claim that if an agent is ignorant of his action, then the agent is not genuinely acting.

Thus, it would be natural to assume that Nietzsche is offering an argument of the following form:

- (1) We are ignorant of (most of) our actions.
- (2) If a person is ignorant of his action, then the action is not under his control.
- (3) Therefore, (most) actions are not under our control.

Given that genuine agency requires that the agent be in control of the action, it would follow that most appearances of agency are illusory.

If this is Nietzsche's argument, though, it runs into a number of problems. First and most obviously, Nietzsche often focuses his criticisms upon Kant's model of agency, yet Kant himself claimed that we can never be certain which maxims we are acting upon.³ If (1) merely amounts to the claim that we cannot be certain about our motives and purposes in acting, then Nietzsche and Kant both accept (1). Kant would also seem to accept some version of (2), yet he certainly denies (3). So, if this is not to be a mere standoff, more must be said.

An attempt to resolve the standoff will lead to another problem. Suppose Nietzsche has convincing arguments establishing that we are typically unaware of some aspects of our actions, and are therefore typically mistaken about what, exactly, we are doing and why. An opponent could accept this point and argue as follows: no one thinks that an agent must be aware of *every aspect* of her action in order to be genuinely in control of it. After all, agency does not require omniscience. Nietzsche would not deny that there are degrees of control: on one end of the spectrum we might put involuntary spasms and reflexes; a bit further along, actions motivated by unconscious attitudes; further still, actions which are influenced by conscious thoughts; and, at the furthest point, perhaps never reached, actions which are completely under the agent's control. Somewhere along this spectrum, there will be a point at which it is false to say that the agent was genuinely acting; somewhere else, there will be a point at which it is true to say that the agent was genuinely acting. But no one, I think, would say that the latter point is at the far end.⁴

Thus, although Nietzsche means his reflections on self-ignorance to constitute some kind of objection to traditional accounts of agency, it is unclear how they would undermine any plausible theory of agency. So we are left with a puzzle: do Nietzsche's reflections on the prevalence of self-ignorance have any interesting philosophical consequences, or are they simply compatible with most theories of agency? To answer this question, we need to investigate the particular forms that self-ignorance takes, as well as the role that conscious thought plays in the causal genesis of our actions. To do so, we need to examine Nietzsche's drive psychology.

II. NIETZSCHE'S DRIVE PSYCHOLOGY

Nietzsche's principal explanatory token is the drive (*Trieb, Instinkt*).^{5,6} In explaining an action, a person's character, a practice, a custom, a morality, a value, a way of life, Nietzsche will appeal to the presence of certain drives.

Nietzsche typically invokes drives to explain broad patterns of behavior rather than particular instances of action. When Nietzsche is explaining particular actions, he will more often appeal to desires, impulses, and emotions. Nietzsche frequently uses the term "affect" (*Affekt*) to refer to this class of psychic states (see, for example, GS 1, BGE 23, and GM III 20). However, it is important to note that drives are explanatorily prior to affects. For example, consider an agent who desires to spend the evening with her friends. The desire explains the particular action (spending the evening with friends). However, a drive explains the desire's presence: it is because the agent has a drive toward sociality that the thought of spending an evening with friends appeals to her.

Thus, when a drive explains a particular instance of action, it does so at one remove. The drive triggers affects that then incline the agent to perform particular actions. In the following sections, we will see that drives also trigger particular *thoughts, patterns of attention,* and *evaluative outlooks*. These mental states and events can also be invoked to explain particular actions.

It is illuminating to compare Nietzsche's drive psychology to other accounts of drives. Although eighteenth- and nineteenth-century philosophers frequently appealed to drives to explain human behavior, most limited their appeals to a few distinct, carefully demarcated drives. Kant names three drives: the drive to self-preservation, the drive to reproduction, and the drive to community.⁷ Schiller also counted three fundamental drives:

the thing-drive (*Stofftrieb*), the form-drive (*Formtrieb*), and the play-drive (*Spieltrieb*).⁸

Nietzsche is much more profligate: in *Daybreak* 422 he mentions our “fifty separate drives.”

In fact, he is undercounting: throughout his corpus he names over one hundred distinct drives. To give a brief sampling, these include: the drive to imagine metaphors, to nature, to logic, to rest, to fight, to distinguish oneself, to create art, to avoid boredom, to knowledge, to appearances, to religion, to freedom, to domination, to sociality, to sex, to property, to politics, to play, to lie, to life, to self-preservation, and to truth.⁹ Drives are *to* (*zu*) the most diverse range of ends.

Moreover, Nietzsche always envisions drives as deeply interconnected: drives do not exist as isolated tendencies, but are intertwined with many other drives. For example, the drive to dominate can be intertwined with the ascetic drive, the drive to knowledge, the drive to religion, and so forth. This is why Nietzsche repeatedly speaks of the *chemistry* of drives: chemistry studies discrete entities that interact, combine, and break up to form new complexes.¹⁰ So if we pick an end which human beings seem to pursue – say, self-preservation – and look closely, we will find not one drive, but many drives, interacting to produce the tendency toward the end. This is why Nietzsche provides no definite list of drives.

Thus, Nietzsche departs from traditional accounts of drives in two ways: his drives are indefinitely multiple, and deeply interconnected.

III. DRIVES AND VALUES

We have noted that drives are motivational tendencies explaining broad tracts of behavior, and we have seen that Nietzsche countenances indefinitely many deeply

interconnected drives. But Nietzsche doesn't simply claim that drives explain behavior; he also makes two claims that are, initially, much more mysterious. First, he claims that drives "interpret the world" (WP 481). Second, he claims that drives include an evaluative component (WP 260). In fact, he thinks that these two points are actually the very same point: "Every center of force adopts a *perspective* toward the whole *remainder*, i.e., its wholly determinate *valuation*..." (WP 567). What are we to make of this?

Here it will help to begin with an obvious point: drives are *embodied* in organisms; drives do not exist in isolation. So, we can ask how a drive, *considered as part of an organism*, might be said to evaluate or interpret the world.

Nietzsche's point, I believe, is this: organisms' perceptions of the world are both *partial* and *evaluatively loaded*. The presence of a drive explains the particular manner in which an organism's perceptions are partial and evaluatively loaded.

We can begin with the claim that perceptions are partial, turning to the claim that perceptions are evaluatively loaded in Section 5. The idea is that anyone contemplating or experiencing a situation will be entertaining some description of the facts of the case; this description will be partial, couched in these terms when it might have been couched in those terms. Moreover, this description will influence, constrain, and sometimes even determine the way that the person goes on. This shows up even in relatively simple cases: until you stop seeing the geometrical problem in *this* way, you won't be able to solve it. It is also obvious in the case of discussion: often, changing someone's mind is achieved by getting him to use different descriptions of the same facts ("Don't think of it as stealing, think of it as taking a little something from an incredibly rich, greedy corporation"; "Don't think of it as making a joke, think of it as hurting Tom's feelings").

Stuart Hampshire makes this point quite clearly:

It is misleading to speak of 'the facts of the situation' in such a way as to suggest that there must be a closed set of propositions which, once established, precisely determine the situation. The situations in which we must act or abstain from acting, are 'open' in the sense that they cannot be uniquely described and finally circumscribed. Situations do not present themselves with their labels attached to them... (Hampshire 1949, 476)

In a footnote to the above passage, he adds:

The word 'fact', here as always, is treacherous, involving the old confusion between the actual situation and the description of it; the situation is given, but not 'the facts of the situation'; to state the facts is to analyze and interpret the situation. And just this is the characteristic difficulty of actual practical decisions, which disappears in the text-book cases, where the 'relevant facts' are pre-selected. (Hampshire 1949, 476)

Descriptions of situations, in virtue of the fact that they characterize the situation in a determinate way, involve simplification and incompleteness. Certain details are emphasized at the expense of others. In this sense, the descriptions are inescapably partial.

Nietzsche thinks of this phenomenon not in terms of differing *descriptions* of the situation confronting us, but as differing *views* of the situation. Speaking of views instead of descriptions emphasizes the *immediacy* and *intractability* of these ways of looking at the situation. Descriptions or views are immediate in that they are not, in general, the product of conscious choice (though of course they can be, and when they are the person may be engaged in an activity of great moral import). Descriptions or views are intractable in that it takes work, often hard work, to revise them.

Nietzsche's claim that drives evaluate or interpret the world is best interpreted, I believe, as the claim that drives engender selective views of the world. We can address this claim in two stages. First, drives and affects can influence the *salience* of features of the environment. Second, drives and affects can produce a view that *distorts* its object.

IV. SALIENCE, SELECTIVITY, AND DISTORTION

Consider the way in which having an end or a purpose can influence one's experience of the world. I have an appointment with Pierre at four o'clock; delayed by an unexpected event, I enter the café a bit late, wondering if he will still be there. I look around, and the *faces* of the patrons leap out at me, impress themselves upon me, each one showing itself not to be Pierre. The café itself, the furniture, the waiters, the food, the coffee machines, all are peripheral, scarcely noticed. The faces are salient in a way that these things are not.¹¹

Here my end of finding Pierre influences the salience of certain features of my situation. In fact, it is easy to see this even with the manifestations of simple instincts, such as hunger. When I am hungry, the presence of food is salient; walking down the street, I notice each restaurant, I see each piece of food. When I am not hungry, the presence of food will often be mere background, barely noticed.

In the above examples, having an end or experiencing an affect influences the salience of certain features of the environment. However, there are also more dramatic forms of influence. Nietzsche believes that drives and affects structure the way in which the situation at hand is presented to us. This is easiest to see in cases of powerful emotions, such as hatred. Consider what it is to hate someone: it is to find everything about him

despicable, it is to focus on all of his flaws and ignore all of his virtues. To hate someone is, in part, to be blind to the good in him and to focus exclusively on the bad. It is to bear a certain selective orientation toward him, to find certain features (the despicable ones) salient and others (the good ones) peripheral.

In the aforementioned respects, hatred resembles hunger, in that both states generate perceptual saliences. But the effects of hatred are not limited to the generation of saliences. In many cases, hatred not only leads the agent to focus on certain features at the expense of others, but also *distorts* its object.

This point is more difficult, and in order to analyze it, it will be helpful to have a detailed and concrete example before us. In a famous passage, Iris Murdoch provides an excellent illustration of this phenomenon, and in order to preserve the richness of her example I will quote it at length:

A mother, whom I shall call M, feels hostility to her daughter-in-law, whom I shall call D. M finds D quite a good-hearted girl, but while not exactly common yet certainly unpolished and lacking in dignity and refinement. D is inclined to be pert and familiar, insufficiently ceremonious, brusque, sometimes positively rude, always tiresomely juvenile....

Thus much for M's first thoughts about D. Time passes, and it could be that M settles down with a hardened sense of grievance and a fixed picture of D, imprisoned (if I may use a question-begging word) by the cliché: my poor son has married a silly vulgar girl. However, the M of the example is an intelligent and well-intentioned person, capable of self-criticism, capable of giving careful and just *attention* to an object which confronts her. M tells herself: 'I am old-fashioned and conventional. I may be prejudiced and narrow-minded. I may be snobbish. I am certainly jealous. Let me look again.' Here, I assume that M observes D or at least reflects deliberately about D, until gradually her vision of D alters. If we take D to be now absent or dead this can make it clear that the change is not in D's behavior but in M's mind. D is discovered not to be vulgar but refreshingly simple, not undignified

but spontaneous, not noisy but gay, not tiresomely juvenile but delightfully youthful, and so on.

(Murdoch 1985, 17-18)

Murdoch is careful to emphasize that the *situation itself* remains constant: D's behavior does not change at all. Nevertheless, M's *view of the situation* undergoes dramatic changes. M initially sees D's behavior as vulgar, undignified, and juvenile; later, she sees the same behavior as refreshingly simple, spontaneous, and delightfully youthful. It takes work – hard, self-critical work – to achieve this later view.

This passage reveals the way in which views are selective. Moreover, it illustrates the way in which the agent's attitudes influence her view. It is partly *because* M is jealous that she initially views D as brusque and juvenile.

As Murdoch's example makes clear, while M's jealousy is making certain features of the situation salient, it is also doing something more: it is leading M to experience those features in a *determinate way*. Thus, it's not just that D's gestures are salient, but that the gestures are seen as brusque, whereas later they will be seen as delightfully youthful. (Here it is important to notice that M is not first experiencing a neutral movement of the hand and then interpreting it as brusque; rather, she immediately sees the hand movement as brusque.)

This passage provides an excellent illustration of the phenomenon in which Nietzsche is interested: the way in which affects influence experience. Nietzsche believes that the Murdochian case is perfectly ordinary: affects and drives *always* structure reflective thought. Murdoch presents the case as if M rids herself, once and for all, of an affect that was distorting her view of D. But there is no reason to take this for granted. M's new view of D might be the product of some other affect: perhaps she takes pleasure in thinking of her son as being married to a wonderful woman, and is thereby led to a different view.

Perhaps the new view is *even more* distorting; perhaps D *is* annoying, brusque, an affront to all who come near her. The mere fact that the new view presents D in a more *appealing* light does not entail that the new view presents D more *accurately*.

In short: the new view is as likely to be distorted and biased as the old view.¹² For *every* view is influenced by the agent's drives and affects. This is why Nietzsche derides the quest for "immaculate perception," perception that is not influenced by any affects (Z 2 "On Immaculate Perception"). These influences are particularly clear in the case of extreme emotions, such as hatred and jealousy, but Nietzsche believes that the influence occurs, in more subtle ways, with every attitude. Here he follows Schopenhauer, who claims that "every inclination or disinclination twists, colors, and distorts not merely the judgment but even the original perception of things" (Schopenhauer, *The World as Will and Representation*, vol. II, p. 373).¹³

Thus, Nietzsche believes that drives and affects structure the agent's experience of the world in two ways. First, they generate perceptual saliences, causing certain features of the environment to impress themselves upon the agent while others recede into the periphery. Second, drives and affects *distort* the environmental features, in the sense that the particular way in which the agent experiences a situation is a function of facts about the drives and affects that he harbors.¹⁴ This is why Nietzsche speaks of drives "coloring," "gilding," "lighting," and "staining" the world (see for example GS 7, 139, 152, 301; BGE 186).¹⁵

V. EVALUATIVE ORIENTATIONS

The preceding sections have discussed Nietzsche's claim that our views of our environment are always selective and are often distorting. The particular manner in which the views are selective and distorting is determined, at least in large part, by our affects and our ends. We saw earlier that our *drives* are responsible for generating our affects and ends. So we can say that our views of the environment are a function of our drives.

With this in mind, we can now understand what Nietzsche means when he claims that having a drive is having a valuation. Our views of the world are selective, emphasizing certain features at the expense of others, presenting objects as oriented toward ends of ours, presenting situations in affectively charged ways. This selective orientation can be understood as an *evaluative* orientation. For a person can be said to evaluate or interpret a thing insofar as he bears a certain orientation toward it or reacts to it in a certain way.

Take an example: consider what it means to say that human beings value the sexual. It does not simply mean that human beings tend to have thoughts such as "sex is valuable." Rather, it means that we experience the world in sexual terms: we are attracted to particular situations, we experience certain features of our environment as alluring, we are inclined to pursue certain ends, we dwell on the presence of the sexual. This is why Nietzsche writes,

The extent of moral evaluations: they play a part in almost every sense impression. Our world is *colored* by them. (WP 260)

There is no doubt that all sense perceptions are wholly permeated with value-judgments... (WP 505)

We experience the world as *oriented*, and the orientation is a function of our attitudes and ends. Nietzsche's point is that the world does not present itself as an indifferent array of

inert facts. The world tempts and repulses, threatens and charms; certain features impress themselves upon us, others recede into the periphery, unnoticed. Our experience of the world is fundamentally value-laden.¹⁶

VI. DRIVE PSYCHOLOGY, SELF-IGNORANCE, AND AGENCY

We have seen that drives generate evaluative outlooks. But recall that our purpose, in investigating Nietzsche's drive psychology, was to determine how this psychology bears on his claims about self-ignorance. In particular, we wanted to determine what forms of self-ignorance Nietzsche is concerned with, and how they might threaten philosophical models of agency. We will now turn to this topic.

Nietzsche's drive psychology has several implications. Drives are indefinitely multiple and deeply intertwined, and the desires, emotions, and other affects which prompt particular actions are products of these intertwined drives. Accordingly, actions have highly complex etiologies, and typically subserve a multitude of ends. Given the complexity of action's etiology, it is to be expected that agents will fail to recognize many of the drives and affects which motivate the action. Thus, Nietzsche writes,

However far a man may go in self-knowledge, nothing however can be more incomplete than his image of the totality of drives which constitute his being. He can scarcely name even the cruder ones: their number and strength, their ebb and flood, their play and counterplay among one another, and above all the laws of their nutriment remain wholly unknown to him. (D 119)

Here, Nietzsche notes that agents are typically ignorant of factors including (i) which drives and affects figure in the causal history of the action, (ii) how these drives and affects interact, and (iii) how these drives and affects manifest themselves.¹⁷

Even when agents do grasp particular drives and affects, they often have an incomplete understanding of both their natures and their relations to other elements of the mental economy. Moreover, drives influence the agent's view of the world. The drives influence both the *salience* of particular features of the environment and the *content* of our perceptual experiences. Agents typically fail to recognize these influences and their effects.

The complexity of action creates a problem for traditional models of agency. In the tradition, many philosophers have linked genuine agency to the capacity for *reflection*. It is often thought that a person who is capable of reflecting on her motives and purposes, and whose reflection is capable of having a causal impact on her actions, is a paradigm of agency. For example, Locke writes that the mind has “a power to *suspend* the execution and satisfaction of any of its desires.” The mind can “consider the objects of [these desires]; examine them on all sides and weigh them with others. In this lies the liberty that man has” (John Locke, *An Essay Concerning Human Understanding*, 263). That is, freedom is the capacity to reflect on one's desires and act as one deems appropriate.

Nietzsche's drive psychology complicates this picture. Suppose we thought, with Locke, that reflection enabled the agent to “suspend” the causal impact of drives and affects and to examine them “on all sides.” We would then envision reflection as enabling the agent to make an unbiased assessment of these drives and affects; free of their influence, the agent could consider them “on all sides.” To be sure, the agent would sometimes fail: his reflective decisions would sometimes be overpowered by recalcitrant desire, as when he decides to stop drinking at the party and finds himself succumbing to the temptation for yet

another drink. But in these kinds of cases, the problem would be that the agent simply had insufficient power to enforce compliance with the edicts of reflection. Reflection followed by decision and action would remain the exemplary case of agency.

Yet if Nietzsche's drive psychology is correct, then this picture is misguided. The problem can be illustrated by returning to Murdoch's example. Envision M at a somewhat earlier stage of her relationship with D, before M investigates the connections between her jealousy and her perceptions of D. Imagine that M needs to decide how to act toward D. M reflects, sees that D is vulgar and brusque, and therefore decides to spurn her. But notice that the perception of D's behavior as vulgar and brusque is, in part, a result of M's jealousy. So M's reflective choice, M's assessment of potential actions in light of evidence from observation, is influenced by M's jealousy.

Murdoch's example makes it clear that the traditional model offers an impoverished and flawed conception of what happens when agents deliberate. Reflection does not enable the agent to escape from, much less to "suspend", the influences of drives and affects. For the very data upon which one reflects will be influenced, and sometimes distorted, by the drives and affects from which one thinks one is stepping back.

Once we recognize the complexity of the deliberative process, as disclosed by Nietzsche's drive psychology, the central idea behind the traditional model of agency looks implausible. This is the claim that control is ensured by the fact that the agent deliberates and acts in accordance with her deliberation. Consider once again M. In a sense, M controls her action: she reflects, makes a decision, and acts in accordance with the decision. But in a deeper sense, her actions are out of her control, for she does not realize that her *jealousy* is moving her in ways that she would disavow, were she aware of them. The nature of M's decision is influenced: M's conscious thought is "secretly guided and channeled" by her

jealousy (BGE 3). She acts in accordance with her choice, but there is good reason to say that her *choice* is not under her control. What looks like a paradigmatic case of agency is in fact compulsion by unrecognized attitudes.

Control is not guaranteed by the mere fact that the agent engages in an episode of reflective choice and acts in accordance with the choice, for reflective choice will often be influenced by factors of which the agent is unaware, and will subserve ends which she would reject.

VII. CONCLUSION

We began with a question: how do Nietzsche's claims about self-ignorance bear on our notion of agency? We now have one answer: Nietzsche's reflections establish that appearances of control are often illusory. Once we recognize the extent and nature of self-ignorance, we see that reflective choices – those paradigms of conscious acts, which are typically treated as manifestations of human freedom – in fact take place within a complex of largely unknown and unnoticed biases. Our conscious choices are “secretly guided and channeled” by a variety of unknown factors (BGE 3). It would be a mistake, therefore, to claim that control is ensured by the mere fact that one deliberates and acts in accordance with the results of one's deliberation.

Is Nietzsche arguing, then, that genuine agency is unachievable, an impossible idea? Nothing we have seen compels us to accept that conclusion. Nietzsche has argued that control is threatened by the prevalence of self-ignorance. But control need not be treated as an on/off notion – we might argue that it comes in degrees. As Nietzsche notes, “self-regulation is not something attained immediately” (KSA ix 11[130]). But perhaps it can be

attained, if only in a gradual and piecemeal fashion. Nietzsche does, after all, repeatedly speak of “self-determination,” “tak[ing] responsibility for oneself,” and being a “sovereign individual” (HH P 3, TI “Skirmishes” 38, GM II 2). This suggests that Nietzsche does have some conception of genuine agency, which contrasts with the degenerate forms of agency manifested by most individuals.

If Nietzsche is offering a conception of genuine agency, what might it involve? If the problem is self-ignorance, the solution may be self-knowledge: we extend our powers of control by acquiring deep and comprehensive self-knowledge.¹⁸ An agent who knows very little about herself manifests less control than an agent with a thorough understanding of her own motivations. Might Nietzsche in this way link genuine agency to self-understanding? If we could see how increasing self-knowledge might be linked to increasing powers of control, if we could do this without gainsaying Nietzsche’s claim that self-knowledge is often a path to ruin (EH “Destiny” 9, WP 492), then I believe we would have before us Nietzsche’s conception of genuine agency. But this is a large and complex undertaking, which must be left for another occasion.^{19,20}

References

- Anscombe, G.E.M. 2000. *Intention* (Cambridge: Harvard University Press).
- Assoun, Paul-Laurent. 2000. *Freud and Nietzsche*, trans. R. Collier (New Brunswick: Athlone Press).
- Bratman, Michael. 2001. "Two Problems about Human Action," *Proceedings of the Aristotelian Society* 101: 309-326.
- Brewer, Talbot. 2002. "Maxims and Virtues," *Philosophical Review* 111: 539-572
- Hampshire, Stuart. 1949. "Fallacies in Moral Philosophy," *Mind* 58: 466-482.
- Katsafanas, Paul. 2005. "Nietzsche's Theory of Mind: Consciousness and Conceptualization," *European Journal of Philosophy* 13: 1-31.
- . Forthcoming. "Nietzsche's Philosophical Psychology," in Richardson and Gemes (forthcoming).
- Locke, John. 1975. *An Essay Concerning Human Understanding* (Oxford: Oxford University Press).
- Murdoch, Iris. 1985. *The Sovereignty of Good* (New York: Routledge).
- Parkes, Graham. 1994. *Composing the Soul* (Chicago: University of Chicago Press).
- Richardson, John. 1996. *Nietzsche's System* (New York: Oxford University Press).
- 2004. *Nietzsche's New Darwinism* (New York: Oxford University Press).
- Richardson, John and Ken Gemes (eds.). forthcoming. *The Oxford Handbook of Nietzsche* (New York: Oxford University Press).
- Sartre, Jean-Paul. 1959. *Being and Nothingness*, trans. Hazel Barnes (New York: Washington Square Press).
- Schiller, Friedrich. 1983. *On the Aesthetic Education of Man*, trans. E. Wilkinson and L.A. Willoughby (New York: Oxford University Press).
- Schopenhauer, Arthur. 1966. *The World as Will and Representation*, 2 volumes, trans. E.F.J. Payne (New York: Dover).
- Stampe, Dennis. 1987. "The Authority of Desire," *Philosophical Review* 96: 335-381.

Notes

¹ When quoting from Nietzsche's works I use the translations by Kaufmann and Hollingdale, though I have sometimes made minor modifications for the sake of clarity.

² The locus classicus for this view is Anscombe (2000). Anscombe argues that intentional actions are distinguished by the fact that the agent stands in a distinctive epistemic relation to them. She labels the requisite epistemic relation "non-observational knowledge": if the agent is acting intentionally, then (a) he knows what he is doing, and (b) this knowledge is not based upon observation.

³ See, for example, *Groundwork of the Metaphysics of Morals* 4:407.

⁴ In fact, the position of these points is likely to be indefinite and context-dependent.

⁵ Nietzsche seems to regard *Trieb* and *Instinkt* merely as terminological variants; he will sometimes alternate between the two in the same sentence (see, for example, GS 1). Here, I will simply use the term *drive* to translate both *Trieb* and *Instinkt*. (I use *drive* instead of *instinct* because, as we will soon see, the English term "instinct" has misleading connotations.)

⁶ Given the prominence of drives in Nietzsche's writings, one would expect to find a large literature analyzing Nietzschean drives. Surprisingly, this is not so; sustained treatments of the topic are relatively rare. There are a few major exceptions: Assoun (2000), Parkes (1994), Richardson (1996) and (2004). I offer an extended discussion of drives in Katsafanas (forthcoming).

⁷ *Religion within the Limits of Mere Reason*, 6:615.

⁸ See especially Letter 14 of *On the Aesthetic Education of Man*.

⁹ For these drives, see: KSA i 887, D 113, KSA vii 16[13], UM II 6, HH 281, AOM 211, WS 31, TI "Skirmishes" 39, KSA vii 3[64], D 143.

¹⁰ Cf. Assoun, 70-82.

¹¹ I am of course alluding to Sartre's discussion of Pierre in *Being and Nothingness*. See pages 40-42 of the H. Barnes translation.

¹² This point needs to be treated delicately; it can easily be misinterpreted. The claim that drives "distort" perceptual experiences should not be taken to imply that there is some pure, undistorted case of perceptual experience. On the contrary, Nietzsche is careful to emphasize that the way to achieve a less partial view of a state of affairs is not to purge the influences of affects from one's view – that would be impossible – but to alternate between many affectively-charged views of the state of affairs: "The task: to *see things as they are!* *The means:* to be able to see with a hundred eyes, from *many* persons! It was a mistake to emphasize the impersonal and to characterize seeing with the eyes of one's neighbor as moral. To see from *many* neighbors and with purely personal eyes—that is the right thing.

The “impersonal” is merely the personal *weakened*, something feeble: it can admittedly be useful every now and then, where there is a need for dispelling the clouding of passion from one’s vision” (KSA ix 11[65]). Cf. D119, 432, 539; GS 301; GM II 12; BGE 230; CW Epilogue; WP 259, 481, 567, 636, 643, 804. For a useful discussion of this point, see Richardson (1996, 35-39).

¹³ Schopenhauer also discusses this idea throughout Volume II, Chapter XIX. For contemporary discussions of related ideas, see, for example, Stampe (1987) and Brewer (2002).

¹⁴ The general idea here is clear enough: drives manifest themselves in part by influencing the content of perceptual experience. However, a number of difficulties and complications arise when we attempt to analyze this idea more precisely. I address this point in much greater detail in Katsafanas (forthcoming).

¹⁵ Elsewhere, I have examined Nietzsche’s remarks on consciousness. I believe Nietzsche identifies conscious mental states with states that have *conceptually articulated content*, and unconscious states with states that have *nonconceptually articulated content*. It takes a good deal of work to explicate these notions, and I will not attempt a summary here. However, it is worth noting that, if this interpretation of Nietzsche’s view of consciousness is correct, then Nietzsche’s claims about the distorting effects of conscious perceptual experience would be readily comprehensible. For a discussion of this point, see Katsafanas (2005).

¹⁶ To avert a potential misunderstanding, note that none of this needs to be conscious. An organism need not, and typically will not, be aware of the evaluative outlook manifested in its orientation toward its environment. This is clearly true of non-self-conscious animals: a bee may perceive the world as oriented toward the securing of pollen – the flowers and potential predators may be salient in a way that the rocks and clouds are not – but the bee of course lacks the capacity to become aware of the way in which its view is selectively oriented. A self-conscious animal, such as a human being, *can* become aware of the partiality and selectivity in its orientation, but it *needn’t*.

¹⁷ I take it that when Nietzsche mentions the drives’ laws of “nutriment,” he is referring to the way in which drives discharge or manifest themselves.

¹⁸ Of course, this process could never be complete: our actions are endlessly complex, so there is no point at which we could truthfully say that there is nothing more to learn about them.

¹⁹ I address these topics in Katsafanas (forthcoming) and work in progress.

²⁰ I thank the audience at the APA Central Division meeting for their comments. I am especially grateful to Lanier Anderson, for a long and detailed discussion of the ideas in this paper. Thanks also to Danielle Slevens for her helpful written comments.