



The Oxford Handbook of Nietzsche

John Richardson (ed.), Ken Gemes (ed.)

<https://doi.org/10.1093/oxfordhb/9780199534647.001.0001>

Published: 2013

Online ISBN: 9780191749803

Print ISBN: 9780199534647

CHAPTER

31 Nietzsche's Philosophical Psychology

Paul Katsafanas

<https://doi.org/10.1093/oxfordhb/9780199534647.013.0032> Pages 727–755

Published: 16 December 2013

Abstract

This article examines Nietzsche's notion of a drive. It addresses three main questions: first, what is a drive? Second, what type of awareness do we have when we are being moved by a drive? Third, what is the relationship between being moved by a drive and reflectively choosing to perform an action? A drive is a disposition that lead agents to evaluative orientation. Drives manifest themselves by structuring the agent's perceptions, affects, and reflective thought. Drives do not simply arise in response to external stimuli; they actively seek opportunities for expression, sometimes distorting the agent's perception of the environment in order to incline the agent to act in ways that give the drives expression.

Keywords: [drive](#), [human action](#), [morality](#), [action](#), [disposition](#), [Friedrich Nietzsche](#)

Subject: [History of Western Philosophy](#), [Moral Philosophy](#), [Philosophy of Mind](#), [Philosophy](#)

Series: [Oxford Handbooks](#)

Collection: [Oxford Handbooks Online](#)

Freud claimed that the concept of *drive* is “at once the most important and the most obscure element of psychological research” (“Beyond the Pleasure Principle,” 1957: vol. 18, p. 34). It is hard to think of a better proof of Freud's claim than the work of Nietzsche, which provides ample support for the idea that the drive concept is both tremendously important and terribly obscure.

Nietzsche tells us that psychology is “the path to the fundamental problems” (BGE 23). Included among these “fundamental problems” are the nature of agency, freedom, selfhood, morality, and evaluation. The psychological concept that is the key to these notions, Nietzsche's principal explanatory token within psychology, is the *drive* (*Trieb*, *Instinkt*).¹ For example, Nietzsche tells us that the self is a relation of drives (BGE 6, 9, 12), and he claims that willing should be understood in terms of the operations of drives (BGE 19). If we are to understand these central elements of Nietzsche's thought, we will need an account of his concept of drive.

However, it is far from clear what exactly a drive is. Talk of drives conjures up images of very basic motivational states, such as urges or cravings; it can also bring to mind physiological states. Thus, *The Oxford English Dictionary* tells us that a drive is “any internal mechanism which sets an organism moving or sustains its activity in a certain direction, or causes it to pursue a certain satisfaction...*esp.* one of the recognized physiological tensions or conditions of need, such as hunger and thirst.” Hunger and thirst are indeed what spring to mind when we think of drives. Many commentators assume that Nietzsche has the same understanding of drives, and consequently treat drives either as simple urges and cravings or as purely physiological states.

But these interpretations cannot be correct. Nietzsche does not identify drives with physiological states or simple causal forces. On the contrary, he explicitly *contrasts* his drive psychology with certain “materialistic” explanations of human behavior (*BGE* 12).² Moreover, he tells us that drives “adopt perspectives,” “interpret the world,” and “evaluate.”³ Clearly, physiological states and urges do not do *that*.

The language of valuing, interpreting, and adopting perspectives is ordinarily used only with regard to *agents*. So Nietzsche sometimes seems to be treating drives as agents-within-agents, homunculi with ends of their own. Some commentators have taken this at face value, interpreting drives as homunculi. For example, Peter Poellner writes that “Nietzsche ultimately treats drives not as attributes of agents (like desires) but as agents themselves” (1995: 174). Yet this proposed interpretation encounters its own set of problems. It is difficult to see how there could be any theoretical advantage in explaining agency and selfhood by appealing to entities that already possess the properties of full-fledged agents and selves. Moreover, it would be rather incongruous for Nietzsche, who so vociferously argues against the superfluous positing of subjects, to multiply the number of subjects beyond measure by splintering each human being into a host of homunculi.

Another puzzle arises when we ask how drives *operate*. How does a drive move a self-conscious organism to act? Nietzsche claims that drives operate beneath the level of consciousness. He argues that we are typically ignorant of both what drives we harbor and how these drives move us (*D* 119). This raises the question of how the influence of drives relates to the workings of reflective thought. Consider an example to which Nietzsche often returns: he claims that Wagner’s development can be understood in terms of one drive’s becoming dominant (*CW* Epilogue). Of course, Wagner himself understood his own actions quite differently. After all, Wagner was engaged in some highly reflective activities: he was composing music, self-consciously attempting to inaugurate a new form of culture, and so on. Presumably, Nietzsche is not suggesting that these self-conscious thoughts bear *no* relation to Wagner’s actions. So there is a puzzle concerning the way in which we reconcile claims about the activities of drives with claims about the agent’s reflective thoughts and choices.

Accordingly, Nietzsche’s drive psychology seems to involve an uneasy and possibly incoherent assembly of claims. Drives appear to be at times physiological states and at other times homunculi; moreover, the drive psychology seems to discount the agent’s self-conscious thoughts and choices in ways that are difficult to understand. Yet it would be decidedly odd if Nietzsche’s principal psychological concept bore such obvious inconsistencies. These are not arcane or deeply hidden inconsistencies of the sort that a philosopher might overlook; the tensions are palpable. Could a “psychologist without equal,” a philosopher who regards psychology as the “path to the fundamental problems,” really be this deeply confused about his foundational psychological concept?⁴

In this essay, I will argue that Nietzsche in fact has a coherent and philosophically fruitful account of drives. In order to explicate this account, I will focus on three central questions: first, what is a drive? Second, what type of awareness do we have when we are being moved by a drive? Third, what is the relationship between being moved by a drive and reflectively choosing to perform an action?

Section 1 surveys existing attempts to answer the first and second questions. I argue that these attempts encounter textual and philosophical difficulties, so we need a new account. Section 2 lays some groundwork for this new account, by examining the history of the drive concept. With this historical backdrop in place, Section 3 offers a new account of the nature of drives and the type of awareness that is present in drive-motivated actions. Section 4 then examines the relationship between reflectively choosing to perform an action and being caused by one's drives to perform an action.

1 Interpretations of Nietzschean drives

1.1 First Interpretive Strategy: Drives as Homunculi

We can start with a simple question: what is a drive? To answer this question, let's consider the types of properties that Nietzsche attributes to drives. Nietzsche frequently claims that drives reason, evaluate, interpret, and adopt perspectives. To cite just two examples:

Anyone who considers the basic drives of man to see to what extent they may have been at play... will find that all of them have done philosophy at some time—and that every single one of them would like only too well to represent just *itself* as the ultimate purpose of existence and the legitimate *master* of all the other drives. For every drive wants to be master—and it attempts to philosophize in *that spirit*. (BGE 6)

It is our needs that interpret the world; our drives and their For and Against. Every drive is a kind of lust to rule; each one has its perspective that it would like to compel all the other drives to accept as a norm. (WP 481)

In the above passages, Nietzsche characterizes drives in agential terms. Philosophizing, representing oneself in a certain way, interpreting, and adopting perspectives are typically understood as activities that are performed by full-fledged agents, not by parts of an agent.

Poellner draws attention to this aspect of Nietzsche's view, writing: "It is sometimes not sufficiently appreciated in the literature that, when it comes to specifying the actual mode of operation or agency of these drives, which he in fact seems to conceive as the ultimate agents, Nietzsche invariably uses intentional-mechanistic terms" (1995: 215). Among these terms are "desiring, interpreting, willing, commanding, and obeying" (1995: 216). Poellner notes that "these terms, in their ordinary meanings, imply the presence of consciousness. Can one be said, for example, to be 'interpreting' a text...unless one is aware of there being a text to be interpreted?" (1995: 215).

This raises an interpretive question: just how literally does Nietzsche intend this language? Does he mean to suggest that drives are agents? ↴

p. 730

Poellner takes the agential language quite literally, interpreting drives as homunculi, or agents-within-agents. Clark and Dudrick endorse a similar interpretation. They point out that Nietzsche speaks of drives "commanding and obeying" other drives, and argue that drives therefore "exhibit agency of a sort" (2009: 265). As they put it, Nietzschean drives are "homunculi" or "proto-persons" (2009: 264).⁵ Similarly, Thiele attributes a robust form of agency to drives, including even the idea that drives have "political relations" with one another (1990: 57). He claims that each drive "has its will to dominate and exploit its competitors... the ruling drive(s) provides its own agenda and worldview...The individual...is a battleground of competing drives, each with its own perspective" (1990: 57–8).

These homuncular readings of drives do have an obvious advantage: they fit quite well with Nietzsche's use of agential language in describing drives. Additionally, Nietzsche sometimes does seem to suggest that drives are agents, as in *WP* 270, where he writes that "the assumption of one single subject is perhaps unnecessary; perhaps it is just as permissible to assume a multiplicity of subjects, whose interaction and struggle is the basis of our thought and consciousness in general."

That said, there are compelling philosophical and textual reasons for rejecting the homuncular reading of drives. First, some proponents of the homuncular view fail to appreciate just how radical their thesis is. Some of these readings attribute to drives properties that imply the presence of *self-consciousness*. For example, Thiele speaks of drives having agendas, perspectives, worldviews, and political relations with other drives, but taken literally this implies that drives are aware of one another, communicate with one another, and reason with one another. Thus, Thiele's interpretation would require that each drive have perceptual capacities, communicative capacities, and reasoning capacities. This seems scarcely conceivable.⁶

A second problem arises when we ask how the homuncular view of drives could have any explanatory power. It is difficult to see how there could be any theoretical advantage in explaining agency and selfhood by appealing to entities that already possess the properties of full-fledged agents and selves. For example, take 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.⁷ ↵

p. 731

This brings us to a more fundamental problem with attributing the homuncular view to Nietzsche: it is hard 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*: 'The Four Great Errors' 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 noted above, Nietzsche argues that once we recognize that the self harbors multiple drives, we must reconceptualize the conscious 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 seems 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.

1.2 Second Interpretive Strategy: Drives as Dispositions

The homuncular interpretations take the agential language that Nietzsche employs when describing drives quite literally. At the other extreme, there are interpretations that ignore or downplay this language, assimilating drives to mere urges. For example, Janaway claims that a drive is simply “a relatively stable tendency to activate behavior of some kind” (2007: 214). Indeed, he suggests that drives may be identical to affects, which “are glossed as inclinations and aversions or fors and againsts” (2007: 214). Thus, “we may wonder whether drives and affects are even properly distinguishable kinds” (2007: 213). With Janaway, we have traveled very far from the idea that drives are self-conscious agents; drives are now described as nothing more than inclinations or tendencies.

This minimalist reading of drives is quite common in the Nietzsche literature. For example, Cox suggests that all of the following terms are roughly analogous: drives, desires, instincts, forces, impulses, and passions (1999: 126–7). Schacht claims that the term “drive” or “instinct” applies “to all firmly established dispositions of any significant degree of specificity, however acquired” (1983: 279–80). Hales and Welshon treat drives as “functional states and dispositions” (2000: 159). Leiter seems to identify drives with urges (2007: 99). These interpretations, which I will call dispositional views, agree in their description of drives as members of familiar psychological categories: drives are simply urges, dispositions, or tendencies.

An advantage of the dispositional interpretation is that it renders drives philosophically unproblematic, thereby avoiding the difficulties that plague the homuncular view. However, the dispositional interpretation faces significant problems of its own. First and most obviously, many of these views offer no real explanation of the agential language that Nietzsche uses when he appeals to drives. For example, if Schacht is correct in claiming that a drive is simply a firmly established disposition, what can it mean to say that drives evaluate and interpret? Suppose I have a firmly established disposition to scratch my head when I am thinking; on Schacht’s view, this should count as a drive. But in what sense could this disposition to scratch be said to evaluate or interpret? Or suppose I am firmly disposed to forget my keys every morning. Can this disposition to forget my keys be regarded as adopting a perspective? The questions seem almost nonsensical: the answer seems to be an obvious no. If the dispositional view is to succeed, it will need to explain how drives are appropriate candidates for agential language.

The philosopher who has done the most to address this issue is John Richardson. Richardson emphasizes that Nietzsche employs agential language in describing drives, but Richardson seeks an interpretation of this language that does not require drives themselves to be conscious agents: “when [Nietzsche] says that a drive ‘aims’ at certain ends, ‘views’ the world in a consequent way, and ‘experiences’ certain values within it, none of this is supposed to entail that the drive is conscious” (1996: 38).⁸ Rather, “a Nietzschean drive is a disposition that was selected for a certain result; this result is its individuating goal, which explains its presence and its character” (2004: 39). Drives are simply a certain sort of disposition. Accordingly, Richardson endeavors to make sense of the way in which a disposition can be an appropriate candidate for agential language. Focusing on Nietzsche’s claim that drives *evaluate*, Richardson argues that we can identify values with the ends at which drives aim: “a drive’s values are precisely the goals it drives towards” (2004: 13).

I think Richardson’s approach is illuminating: we should ask whether Nietzsche’s agential language can be applied to dispositions. However, the connection that Richardson draws between being disposed, as a result of selection, toward some end E and valuing E does not seem fully convincing. There are cases in which values and selected dispositions appear to diverge. For example, a typical ascetic who regards sexual activity as disvaluable will nonetheless be strongly disposed, as a result of natural selection, to engage in sexual activity.⁹ Despite the fact that the agent is strongly disposed toward sexual activity, we would typically say that the agent does not value sexual activity. There are also cases of the opposite sort, in which the agent

regards an end E as valuable, but is not disposed toward E. For example, the aforementioned ascetic would view celibacy as valuable, but would be strongly disposed, as a result of selection, not to be celibate.

p. 733 As these examples indicate, being disposed as a result of selection toward an end E and valuing E can come apart. So the identification of values with selected dispositions seems ↴ problematic.¹⁰ Perhaps, though, we can tie valuing to having a *specific kind* of disposition. In an earlier work, Richardson makes a suggestive comment:

Value lies in the way the world is ‘polarized’ for each will and not in any theories or beliefs about value. It lies in how things ‘matter’ to the will and so depends on that deep receptiveness of will that Nietzsche calls ‘affect’ [*Affekt*] or ‘feeling’ [*Gefühl*]. (Richardson 1996: 37)

Here, Richardson suggests that valuing an end E isn’t simply being disposed to E; in addition, valuing E involves having certain affects or feelings. Although Richardson doesn’t pursue the suggestion at length, I think it is the key to unraveling Nietzsche’s remarks about drives. In sections 2 and 3, I will explore this point in detail, arguing that Nietzschean drives are dispositions that induce affective orientations in the agent. Moreover, I will argue that these affective orientations can be understood as *evaluative* orientations.

1.3 Drives and Self-awareness in Action

Before continuing our analysis of the nature of drives, we will need to gain clarity on another aspect of the drive psychology: the way in which drives cause agents to act. Seeing how drives *operate* will help us to understand what drives *are*. Accordingly, in this section I will examine Nietzsche’s characterization of the type of awareness that is present in drive-motivated actions.

When Nietzsche discusses drives, he often emphasizes that agents are ignorant of the way in which drives move them.

p. 734 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) ↴

For this reason, Nietzsche claims that “actions are *never* what they appear to be...all actions are essentially unknown” (*D* 116). But puzzles arise when we ask what Nietzsche means by these claims about self-ignorance in action.

Commentators often interpret Nietzsche as arguing that our actions can proceed independently of conscious monitoring and deliberation. For example, Schacht interprets Nietzsche in this way, illustrating his point with an example of a pianist. Schacht points out that a *novice* pianist may need to consciously attend to his activity, focusing on the positions of the keys, keeping in mind the notes that he wants to play, consciously monitoring his performance, and so on. An *expert* pianist, by contrast, is able “to dispense with the mediation of conscious deliberation and reckoning at each step of the way” (1983: 281). The expert can simply play, without needing to deliberate or consciously attend to his movements. This example of skilled action draws attention to the fact that once we acquire a disposition to perform some activity *A-ing*, we can *A* without deliberating on or attending to our *A-ing*. According to Schacht, Nietzsche is claiming that *all* of our drive-motivated actions have an analogous form, proceeding independently of conscious monitoring.

Now, it certainly is true that many actions occur without conscious monitoring. However, this cannot be the full point of Nietzsche’s claim that agents are ignorant of drive-motivated actions. After all, the piano player is not ignorant of his action in any strong sense: he can attend to his movements at any moment,

without difficulty. Even when he is not explicitly attending to his movements, he certainly *knows* that he is playing the piano. In fact, his playing is exactly analogous to everyday actions such as walking: when I walk to my office, I rarely attend to or reflect on the movements of my legs; indeed, I often walk about in a kind of daze, thoughts occupied with other matters. Nevertheless, I know that I am walking, and my walking is an intentional action.¹¹ If this humdrum type of inattentiveness were all that Nietzsche had in mind when he claims that “all actions are essentially unknown” (*D* 116), then he would be grossly exaggerating a familiar, uncontroversial feature of action.

Moreover, Nietzsche claims that even paradigmatically self-conscious actions are in some sense unknown to us. We should distinguish two claims:

- (1) An agent can perform an action *A* without self-consciously attending to her *A*-ing.
- (2) An agent who *does* self-consciously attend to her *A*-ing can in some sense remain ignorant of her *A*-ing.

Schacht’s analysis illustrates (1), yet Nietzsche more often focuses upon (2). For example, Nietzsche writes, “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). Here, Nietzsche is not claiming that we can act without monitoring our act; he is claiming that even if we do monitor our act, we will in some sense be ignorant of it. Schacht’s point about the dispensability of conscious monitoring seems unable to account for this aspect of Nietzsche’s view. ↵

p. 735

Consider, then, an alternative interpretation of Nietzsche’s remarks on self-ignorance in action: perhaps Nietzsche is arguing that we cannot know our true *motives* for action. As Leiter puts it, “we do not have epistemic access to what the causally effective motives really are” (2002: 104). This interpretation fits the texts somewhat better: passages such as *BGE* 32 and *D* 116, quoted above, certainly suggest that we are mistaken about our true motives.

However, I think this interpretation also falls short of capturing the full truth. While Nietzsche does claim that we are often mistaken about our causally effective motives, this can hardly be the centerpiece of his analysis of reflective agency. The claim that we lack epistemic access to our causally effective motives is widely accepted; indeed, one could argue that even in Nietzsche’s day it was a commonplace. After all, even Kant, whose model of agency Nietzsche wanted to attack, emphasized that we can *never* be certain which motives we are acting upon (*Groundwork* 4: 407; see Kant 1998).¹²

So we are left with a problem. While Schacht is certainly correct to claim that conscious monitoring is not a necessary condition for action, and while Leiter is undeniably right in claiming that we are often ignorant of our motives for action, neither of these points is controversial. If these are the only points that Nietzsche makes about conscious awareness in action, then his account is in no way revolutionary.

1.4 Summary

The prior sections have addressed two central questions about drives: what is a drive, and what type of awareness do we have when acting under the influence of a drive?

The first question led to some problems: the homuncular view of drives seems philosophically and textually problematic, whereas the dispositional view of drives has difficulty accounting for Nietzsche’s use of agential language in describing drives. The second question was also puzzling: Nietzsche emphasizes that we are ignorant of our own actions, but it is difficult to find an interpretation of this claim that renders it philosophically significant. If Nietzsche is simply claiming that action does not require attention, no one

will disagree; if he is merely pointing out that we are often mistaken about our motives, then he is belaboring a truism.

2 A Highly Abbreviated History of the Drive Concept

p. 736 We can gain clarity on the questions of what drives are and how drives engender self-ignorance by situating Nietzsche's account in its historical context. Nietzsche's drive psychology did not develop in a vacuum; the concept of instinct or drive was much discussed in the eighteenth and nineteenth centuries. In a debate spanning several generations, a diverse group of scientists, philosophers, and theologians attempted to explain what instincts ↵ are, how they arise, and how they move organisms.¹³ Here I will examine just one aspect of this debate: the question of how the unreflective nature of drives should be understood. Answering this question will enable us to illuminate Nietzsche's view of drives.

During the eighteenth and nineteenth centuries, the concept of instinct was typically contrasted with the concept of *learned behavior*. There are marvelous examples of the distinction between learned and instinctive behavior in the animal kingdom, many of which fascinated the thinkers of this time. Reimarus, in his *Allgemeine Betrachtungen über die Triebe der Thiere* (1760), draws our attention to the caterpillar, which weaves its elaborate cocoon without having witnessed anything similar. Henry Lord Brougham discusses a species of solitary wasp that gathers grubs and stores them beside its eggs, then departs before the eggs hatch. The grubs serve as food for the larvae that will hatch from the eggs, but the wasp cannot possibly know this. For "this wasp never saw an egg produce a worm [i.e., a larva]—nor ever saw a worm—nay, is to be dead long before the worm can be in existence—and moreover she never has in any way tasted or used these grubs, or used the hole she made, except for the prospective benefit of the unknown worm she will never see" (Brougham, *Dissertations on Subjects of Science concerned with Natural Theology* (1839): I: 17–18; quoted in Richards 1987: 136). These highly complex behaviors are directed at an end of which the animal simply cannot be cognizant.

These complex, unlearned behaviors are attributed to instincts. Thus, in an early treatise on the notion of instinct, we read of Frédéric Cuvier's distinction between instinct and intelligence:

The wolf and the fox who recognize the traps in which they have been caught, and who avoid them, the dog and the horse, who understand the meaning of several of our words and who obey us, thereby show *intelligence*. The dog who hides the remains of his dinner, the bee who constructs his cell, the bird who builds his nest, act only from *instinct*. (Flourens, *Analytical Summary of the Observations of Frédéric Cuvier* (1839); quoted in Proudhon 1994)

Charles Darwin concurs:

An action, which we ourselves require experience to enable us to perform, when performed by an animal, more especially by a very young one, without experience, and when performed by many individuals in the same way, without their knowing for what purpose it is performed, is usually said to be instinctive. (Darwin 1993: 317–18)

p. 737 So the writers of this time period operate with the following dichotomy: some animal behaviors are learned, and therefore require the animal to have awareness of the goal at which the behavior is directed; other behaviors, the instinctive ones, are not learned, and the animal ↵ performing these behaviors lacks awareness of the goal it is pursuing.¹⁴ Thus, the central characteristic of instinctual behavior is that it is in some sense *unknown* or *unreflective*.

Although thinkers of this time agree that instinctual behavior is unreflective, they disagree about what this means. Some thinkers advance a very strong thesis about the lack of awareness in instinct. Consider movements that, though they look purposive, are mere mechanism. A clock is set up so that it ticks away the hours; a car's engine is set up so that it produces movement when stimulated by a depressed pedal. Of course, the clock and car do not in any sense know what they are doing. We might think that organisms are exactly analogous: instincts operate in a purely mechanical fashion, with stimulus S causing behavior B.

Schopenhauer believed that some instincts operate in this way. At several points throughout his work, he compares the animal acting instinctively with the sleepwalker: he writes that instinctive actions have "a remarkable similarity to those of somnambulists" and claims that "insects are to a certain extent natural somnambulists" (*WWR* II: 344). And of course the Cartesians, as well as some German Materialists of the nineteenth century, claimed that non-human animals were mere mechanisms: Descartes writes that "the actions of beasts are similar only to those which we perform without the help of our minds" (Letter to More (1649); quoted in Huxley, "On the hypothesis that animals are automata, and its history" (1874)) and his followers notoriously compared the screams of an animal to the ringing of a bell.¹⁵

These thinkers suggest that the animal acting instinctively is completely unaware of its action. So we have a very strong claim about the unreflective character of instinctive actions:

(1) If an organism instinctively A-s, then the organism is not aware¹⁶ that it is A-ing.

Claim (1) seems accurate with respect to certain organisms. For example, it is hard to imagine that an amoeba oozing toward its prey is doing anything more than acting mechanically, in response to determinate stimuli; it is not as if there can be mediation by thought here. But some writers argue that (1) mischaracterizes the nature of instinctive actions in more complex animals. For example, imagine a wolf that is instinctively hunting a moose. It is difficult to imagine the wolf pursuing the moose, tracking scents, coordinating with other members of the pack, and so on, all the while being ignorant of its actions. It is more plausible to assume that the wolf has some rudimentary awareness of its actions, which enables it intelligently to adjust the means to the fulfillment of its instincts.

How might this work? To employ a somewhat anachronistic source, consider William James, who writes:

We may conclude that, to the animal which obeys it, every impulse and every step of every instinct shines with its own sufficient light...What voluptuous thrill may not shake a fly, when she at last discovers the one particular leaf, or carrion, or bit of dung, that out of all the world can stimulate ↳ her ovipositor to its discharge? Does not the discharge seem to her the only fitting thing? And need she care or know anything about the future maggot and its food? (James 1890: vol. II, pp. 387–8)

p. 738

Or, to choose an example from a book that was in Nietzsche's personal library: Schneider, in *Der Thierische Wille* (1880), writes, "it might easily appear" that the cuckoo "acted with full consciousness of the purpose" when it laid its eggs in another bird's nest. But no: "the cuckoo is simply excited by the perception of quite determinate sorts of nest, which already contain eggs, to drop her own into them, and throw the others out, because this perception is a direct stimulus to these acts. It is impossible that she should have any notion of the other bird coming and sitting on her egg" (quoted in James 1890: vol. II, p. 389). These quotations suggest that instincts operate by presenting the animal with a compelling motive to act in a certain way: the fly experiences a voluptuous thrill in the presence of a bit of dung; the cuckoo is excited by the perception of a certain kind of nest.

In short, an instinct might operate purely mechanically, by producing a series of behaviors; or it might operate at one remove, by producing internal states, such as emotions, desires, and urges, which then

strongly dispose the organism to pursue some end. The animal acting on these internal states may be aware of its progress toward the nest, its pursuit of its prey, and so forth. But it remains ignorant of something else: the purpose of the action, or the ultimate *end* at which its action is directed. For example, the cuckoo knows that it is laying eggs in a nest, but does not know that it is doing so in order that another bird might care for its young. Or, the female wasp knows that it is collecting grubs, but does not know that it is doing so in order to provide food for its future offspring. So we have a second characterization of the unreflective character of instinct:

(2) If an organism instinctively A-s in order to G, then the organism may know that it is A-ing, but does not know that it is A-ing *in order to G*.

To put (2) in a more colloquial form: the organism may know *what* it is doing, but it doesn't know *why* it is doing what it is doing.^{17 18}

p. 739 So we have two different views on the unreflective character of instinct. Instinctive actions might be unreflective in the sense that they involve no awareness whatsoever (1), ↪ or in the sense that they involve no awareness of the ultimate goal of the action (2). We will have to determine which of these views Nietzsche adopts.

2.2 An Interpretive Clue: Schopenhauer on Drives

Before turning to Nietzsche, let's briefly examine one of the greatest influences on Nietzsche's work: Schopenhauer. Schopenhauer discusses drives at some length and endorses a view of type (2). In a wonderful chapter entitled "The Metaphysics of Sexual Love," Schopenhauer examines the workings of the reproductive drive. He claims that the reproductive drive leads human beings to pursue sexual partners, not by blindly impelling them to this end, but by fostering a *distorted orientation* toward the world. The reproductive drive "creates illusions [*Illusionen schafft*]" (WWR II: 566) or a "delusion [*Wahn*]" (WWR II: 541):

Here then, as in the case of all instinct, truth assumes the form of delusion, in order to act on the will. [*Also nimmt hier, wie bei allem Instinkt, die Wahrheit die Gestalt des Wahnes an, um auf den Willen zu wirken.*] It is a voluptuous delusion which leads a man to believe that he will find greater pleasure in the arms of a woman whose beauty appeals to him than in those of any other, or which, exclusively directed to a *particular* individual, firmly convinces him that her possession will afford him boundless happiness...The character of instinct is here so completely present, namely an action as though in accordance with the conception of an end and yet entirely without such a conception, that whoever is urged by that delusion often abhors it and would like to prevent the end, procreation, which alone guides it...(WWR II: 540)

Schopenhauer here reasons as follows. The human reproductive drive aims at reproduction. But when we are in the grip of this drive, we do not believe that we are pursuing reproduction. We believe we are pursuing happiness, or pleasure, or possession of a particular individual. Schopenhauer claims that this belief—or, as he puts it, this *delusion*—is produced by the drive itself. In other words, the reproductive drive manifests itself by leading a person to conceive of his potential sexual partners as supremely alluring, capable of providing him with great happiness and pleasure. The reproductive drive moves us not by generating a blind urge or disposition to copulate, but by producing desires and other emotions, by influencing the way in which the person perceives potential partners, and so on.

Thus, "in all sexual love, instinct holds the reins, and creates illusions [*bei aller Geschlechtsliebe der Instinkt die Zügel führt und Illusionen schafft*]" (WWR II: 566). But the phenomenon is not restricted to the sexual: Schopenhauer believes that all instincts work in this fashion. Accordingly, he claims that animals acting on

instinct “are urged not so much by an objective, correct apprehension, as by subjective representations which stimulate the desire...and that accordingly they are urged by a certain *delusion*...” (WWR II: 541).

p. 740 Schopenhauer holds that drives typically move a person not by blindly impelling him to act, but by structuring his affects, thoughts, and perceptual orientation toward the world.¹⁹ ↵ Crucially, it follows that the agent’s actions—though they may be highly reflective and deliberate, though they may occupy the agent’s attention, though the agent may think of nothing else—are in one sense *unreflective*: the person being moved by the drive is not aware of his ultimate purpose in acting. Thus, the person being moved by his reproductive drive knows that he is pursuing a particular partner; he knows that he is planning a date; he devotes all of his attention to his actions. Yet the deeper purpose of these actions eludes him. While he thinks that he wants his love because she will provide him with immeasurable happiness, Schopenhauer claims that the deeper purpose is less grandiose: reproduction.²⁰

In sum, we can see that Schopenhauer endorses a view of type (2). The claim that a person is being moved by a drive does not entail that the person cannot be acting reflectively, attending to his action, and so forth. Rather, it implies that the agent’s conscious reflection and thought is in the service of a goal of which the agent is ignorant. The drive manifests itself by generating an affective orientation, which then inclines the agent to pursue the drive’s end. So a drive is a disposition that induces an affective orientation.

3 The Nature of Nietzschean Drives

3.1 Drives are Dispositions that Induce Evaluative Orientations

Schopenhauer treats drives as dispositions that induce affective orientations. In this section, I argue that Nietzsche has an exactly analogous understanding of drives. Ultimately, I am going to argue that this account of drives enables us to make sense of Nietzsche’s claim that drives evaluate and interpret. For the affective orientation induced by a drive can be understood as an *evaluative* orientation.

To make sense of these ideas, let’s start with the most obvious way in which having an end or harboring an affect can influence an agent’s view of the world: it can make certain features *salient*. This is easiest to see with the manifestations of simple feelings, such as hunger. When one is hungry, the presence of food is salient: I notice each restaurant, my attention is drawn to each piece of food eaten by passers-by. When I am not hungry, the presence of food recedes: it is often mere background, barely noticed.¹ ↵

With more complex affects and drives, the influences are of course more complex. Hatred is an instructive case. Hating affects perceptual saliences: if you hate someone, you tend to experience everything about him as despicable, focusing on all of his flaws and ignoring all of his virtues. In other words, hatred typically manifests itself by inducing a certain orientation toward the object of hatred: it leads one to find certain features (the despicable ones) salient and others (the redeeming ones) peripheral.

In each of these cases, the affect influences the perceptual saliences, causing certain features to stand out and others to recede into the background.²¹ This is why Nietzsche is concerned with the role of the emotions and other attitudes in deliberation. In deliberation, the presentation of the facts—the selection of some features as salient and others as peripheral—is, at least in part, a function of the attitudes. This is particularly clear in the case of extreme emotions, but Nietzsche believes that it happens in subtler 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” (WWR II: 373).²²

So the first point is that perceptions are selective, and the particular ways in which they are selective is, in part, a function of our drives and affects. But the effects of drives are not limited to selectivity. Nietzsche believes that there is a sense in which drives influence the *content* of experience itself.

Daybreak 119 offers an extended discussion of this phenomenon. Nietzsche starts with a discussion of dreams:

Why was the dream of yesterday full of tenderness and tears, that of the day before yesterday humorous and exuberant, an earlier dream adventurous and involved in a continuous gloomy searching? Why do I in this dream enjoy indescribable joys of music, why do I in another soar and fly with the joy of an eagle up to distant mountain peaks? These inventions, which give scope and discharge to our drives to tenderness or humorousness or adventurousness or to our desire for music and mountains...are interpretations of nervous stimuli we receive while asleep, *very free*, very arbitrary interpretations of the motions of the blood and intestines, of the pressure of the arm and the bedclothes, of the sounds made by church bells, weatherclocks, night-revelers and other things of the kind. That this text, which is in general much the same on one night as on another, is commented upon in such varying ways, that the inventive reasoning faculty *imagines* today a *cause* for the nervous stimuli so very different from the cause it imagined yesterday, though the stimuli are the same: the explanation of this is that today's prompter of the reasoning faculty was different from yesterday's—a different *drive* wanted to gratify itself, to be active, to exercise itself, to refresh itself, to discharge itself...(D 119)

Nietzsche is interested in the fact that the sensory stimuli present from night to night remain relatively constant, while the dreams vary enormously. He attributes the variation in dreams to the activities of different drives: the same sensory stimuli give rise to quite different dreams, depending upon which drives are most active.

The full point of the discussion of dreams is revealed a few lines later:

Waking life does not have this *freedom* of interpretation possessed by the life of dreams, it is less inventive and unbridled—but do I have to add that when we are awake our drives likewise do nothing but interpret nervous stimuli and, according to their requirements, posit their 'causes'? that there is no *essential* difference between waking and dreaming? (D 119)

p. 742

Nietzsche claims that just as drives influence the content of dreams, so too drives influence the content of waking experience. The same sensory stimuli can give rise to quite different perceptual experiences, depending upon which drives are active. This is clearest in the case of dreams; but Nietzsche believes that the same phenomenon occurs, in a more restricted way, in waking life. He provides the following example: "Take some trifling experience. Suppose we were in the market place one day and we noticed someone laughing at us as we went by." He claims that different agents will experience this stimulus in different ways, depending upon which drives are active. Thus, one person will scarcely notice the laughter, another will be angered by it, another will worry over it, another will be led to reflect on the nature of laughter itself, another will be happy. The selfsame stimulus is experienced in quite different ways.

Of course, Nietzsche is not claiming that drives manifest themselves in *exactly* the same way in dreams and in waking life. In dreaming, there is only the slightest connection between sensory stimuli and experience: the sounds of distant clocks might lead to dreams of beautiful music; the murmurs of night-revelers might lead to thoughts of soaring through the air; the entanglement in blankets might lead to dreams of continuous searching. The effects of drives on waking experience are not this dramatic: while the stimulus of laughter can be experienced in a variety of ways—angrily, happily, contemplatively, and so forth—there is clearly less room for creative interpretation than in the case of dreams. So, when Nietzsche says, "there is

no *essential* difference between waking and dreaming,” he does not mean that facts about the world play as little role in determining waking experiences as they do in determining dreams. Rather, he means that in waking, as in dreaming, our experiences are determined not by facts about the world alone, but also by facts about which drives are active. Thus, Nietzsche will speak of affects and drives as “coloring,” “gilding,” “lighting,” and “staining” the world; these terms suggest that affects and drives highlight or even alter aspects of an experience, but not that they *create* the experience in the way that they create dreams (see for example GS 7, 139, 152, 301; BGE 186). Thus, Nietzsche is seeking to undermine the intuitively plausible thought that our perceptual experiences of the world are determined by nothing other than the nature of the world itself.²³

In order to make Nietzsche’s idea more precise, it will be helpful to work with a more detailed example. A famous passage from Iris Murdoch provides an excellent illustration:

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)

p. 743

There are several important features of this example. First, notice that the *situation itself* remains constant: D’s behavior does not change at all. Nevertheless, M’s *view of the situation* changes dramatically. M initially sees D’s behavior as brusque, rude, juvenile, and pert; later, she sees the same behavior as spontaneous, simple, delightfully youthful, and gay. M achieves this latter view by engaging in critical self-assessment, examining the effects of her motives on her perceptions and judgments.

This passage reveals the way in which affects can influence the content of experience: the selfsame situation can be viewed in exceedingly different ways. M’s jealousy not only makes certain features of the situation salient, but also influences the very content of M’s experience. For example, M’s jealousy not only causes D’s hand gestures to be salient; in addition, it leads M to perceive these gestures as juvenile, whereas later they will be perceived as delightfully youthful. In short, the attitude leads the agent to conceptualize the situation in a certain way. (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. In this way, drives and affects influence the content of experience itself.)²⁴

I have quoted this passage at length because it provides a detailed, realistic illustration of the phenomenon in which Nietzsche is interested: the way in which motivational states influence the content of experience. Drives manifest themselves by coloring our view of the world, by generating perceptual saliences, by influencing our emotions and other attitudes, by fostering desires. Thus, Nietzsche’s idea is that the way in which one experiences the world is, in general, determined by one’s drives in a way that one typically does not grasp.²⁵

3.2 This Account of Drives Avoids Problematic Theoretical Commitments and Explains Nietzsche's Use of Evaluative Language

p. 744 This account of drives and affects enables us to make sense of Nietzsche's claim that drives "interpret the world," generate "evaluations," and "adopt perspectives." Section 1.1 pointed out ↵ that these claims have led to some extravagant interpretations of Nietzsche's drive psychology, tempting some commentators to treat Nietzschean drives as homunculi. Evaluation and interpretation are normally understood as highly reflective acts performed by self-conscious beings. I interpret a book or a poem by reflecting on its meaning; I evaluate an action or a trait of character by reflecting on a moral principle. Unless each drive is a self-conscious center of agency, it is difficult to see how drives could do *that*.

However, we can now see that the homuncular view looks appealing only when we have a restricted view of the available options. Poellner and Thiele seem to assume that there are only two options: either drives, considered as isolated entities, have agential properties or they do not. If these were the only two possibilities, the homuncular view would indeed be preferable. After all, Nietzsche certainly does employ agential language with respect to drives, and it stretches the imagination to claim that these are nothing more than colorful metaphors. So the former possibility seems better, despite its air of paradox.

Yet there is another option: we can deny that drives, *considered in isolation*, can reason, evaluate, and interpret, while maintaining that *embodied drives*—drives considered as part of a whole organism—can reason, evaluate, and interpret. Suppose we accept Nietzsche's claim that 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, affectively charged orientation can be understood as an *evaluative* orientation. For example, if Murdoch's M has an immediate view of D as vulgar, brusque, and rude, this view can be understood as constituting a negative evaluation of D.^{26, 27}

Nietzsche often directs our attention to this point, emphasizing the way in which values are manifested in sensory experiences:

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)

p. 745 The visual language in this passage is revealing: Nietzsche's point is that we experience the world in evaluative terms. 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.^{28, 29} ↵

Thus, the link between drives and values is this: drives generate affectively charged, selective responses to the world, which incline the agent to experience situations in evaluative terms. We can summarize this point by saying that drives are dispositions that generate evaluative orientations. Accordingly, Nietzsche writes, "from each of our basic drives there is a different perspectival assessment [*perspektivische Abschätzung*] of all events and experiences" (KSA 12: 1 [58]); in plainer language, each drive generates an evaluative orientation. Thus, we can make sense of Nietzsche's evaluative language without treating drives as homunculi.³⁰

This interpretation also enables us to see what Nietzsche means when he claims that drives induce a form of self-ignorance. Drives influence an agent's behavior by structuring the agent's view of his environment. This structuring has dramatic effects on behavior: consider the difference in the way that M will act prior to and after the change in her view. Or consider Schopenhauer's remarks on sexual love. The lover believes that

he desires his love because she is beautiful, because she will please him like no other, because she will complete his very being. Schopenhauer explains why the lover sees her this way, why he has these thoughts, why he so ardently desires his love, by appealing to a reproductive urge which colors the lover's mental economy. The lover, if Schopenhauer is right, is acting for reasons that he does not grasp. Thus, drives engender self-ignorance in the sense that agents are typically unaware of the way in which their drives direct their thoughts, affects, and perceptions.

3.3 Drives as Psychic Forces

The above interpretation enables us to make sense of both Nietzsche's evaluative language and his claim that drives induce a form of self-ignorance. However, another aspect of Nietzsche's view remains to be explicated. Nietzsche speaks of the "ebb and flood" of our drives, their "play and counterplay among one another," their "growth and nourishment" (*D* 119; *BGE* 6 et passim). This language is familiar to us; drives are almost inevitably associated with active *forces*, *impulsions*, and *pressures seeking discharge*. Although vague and metaphorical, these colorful terms are suggestive; they are capturing something important about the concept of drives. Drives are not simply *responses* to external stimuli; on the contrary, drives seek to manifest themselves. But again, I will argue that this needn't lead us to interpret drives as agents.

We can begin by considering simpler psychic states: desires. Some desires arise as responses to the perception of external stimuli. Walking down the street on a hot summer day, I see an ice cream shop, and this sight creates a desire for ice cream. Walking through the forest, I see a shape lurching out of the trees, and I desire to get away. These desires are affectively charged responses to external stimuli.

But other desires arise in a different way, seeming to have a life of their own. Consider a habitually aggressive, combative person. Part of what it is to be habitually aggressive is to have a recurrent tendency to seek out opportunities for aggression, regardless of whether ↴ the circumstances merit aggressive responses. The aggressive person will typically distort circumstances in order to find these kinds of outlets. He will interpret ordinary, inoffensive behavior as offensive, raging at the driver who cuts in front of him or the cashier who seems distracted. This aggressive tendency consists, in part, in the tendency to see aspects of his environment as *warranting* aggression. Nietzschean drives are supposed to have an analogous form. They have a psychic life of their own: drives do not *await* occasions for expression, but *create* them, by inclining the agent to see certain actions as warranted.

Freud, no doubt influenced by Nietzsche's conception of drives, suggested a similar model of drives. In "Drives and their Vicissitudes," Freud asks what the relation is between the notions *drive* and *stimulus*.³¹ "Stimulus" here serves as Freud's most general term for a motivational state; it is analogous to the contemporary use of terms such as "pro-attitude" or "desire." So Freud is asking whether drives are just desires. He answers with a qualified "yes." Drives can be understood as a type of stimulus (or desire), but if we do group them in this way it is important not to think that all stimuli function in the same way. For there are two differences between drives and other stimuli. First, "a drive stimulus does not arise from the external world but from within the organism itself." Second, many stimuli operate with

a single impact, so that [they] can be disposed of by a single expedient action. A typical instance of this is motor flight from the source of stimulation. These impacts may, of course, be repeated and summated, but that makes no difference to our notion of the process and to the conditions for the removal of the stimulus. A drive, on the other hand, never operates as a force giving a *momentary* impact but always as a *constant* one. (Freud 1957: vol. 14, p. 118)

Summarizing these points, Freud writes that the essential nature of drives is "their origination in sources of stimulation within the organism and their appearance as a constant force" (Freud 1957: vol. 14, p. 119). So

drives have two features: drives do not await external stimuli, but manifest themselves independently of external stimuli; moreover, drives are not momentary occurrences, but are relatively constant.

Start with the second point. Drives need not be constant in the literal sense of being active at each moment; rather, they are constant in the sense that they arise, with some regularity, throughout the individual's life. Hunger provides a good example: although we are not *always* hungry, there is a sense in which hunger is a constant motive. For hunger cannot be eliminated once and for all; it can only be put into abeyance. Likewise, drives cannot be eliminated, but only temporarily sated.

Turn now to the first point. When a drive is active, it leads the agent to engage in behavior that satisfies the drive. The drive does not await appropriate stimuli or occasions for discharge. Again, take hunger. Though hunger is sometimes roused by external stimuli, such as the sight or smell of food, hunger can also arise independently of any external stimuli. Presumably brought about by physiological conditions, hunger can arise at the most inopportune times, and will not slacken until it is, to some extent, satisfied by the acquisition of some object. ↵

p. 747

When Nietzsche writes of drives being active, ebbing and flooding, and seeking discharge, he has something similar in mind. Drives arise independently of external stimuli, and once they have become active, they will seek discharge. The fact that drives are active and do not arise in response to external stimuli creates a problem. In many cases, a drive will be active in conditions that do not provide the agent with appropriate objects with which to satisfy the drive. Just as we can be hungry when there are no opportunities to eat, we can be angry when there are no occasions for anger. For example, suppose the aggressive drive is active in a situation in which the individual has not been threatened or provoked. Nietzsche tells us that the drive will *seek outlets*—seek objects on which to vent itself.

To clarify this point, it will be helpful to draw on a useful distinction that Freud introduces. Freud distinguishes between the *aim* [*Ziel*] and the *object* [*Objekt*] of the drive. The aim of the drive is its characteristic goal, in terms of which it is individuated from other drives. The aim of the sex drive is sexual activity; the aim of the ascetic drive is ascetic activity; and so on. Freud remarks “although the ultimate aim of each drive remains unchangeable, there may yet be different paths leading to the same ultimate aim” (Freud 1957: vol. 14, p. 118). Thus, he introduces the notion of the drive's object.

The object of a drive is the thing in regard to which or through which the drive is able to achieve its aim. It is what is most variable about a drive and is not originally connected with it, but becomes assigned to it only in consequence of being peculiarly fitted to make satisfaction possible...It may be changed any number of times in the course of the vicissitudes which the drive undergoes during its existence...(Freud 1957: vol. 14, p. 118)

The aim of a drive is its characteristic form of activity. The sexual drive aims at sexual activity; the aggressive drive aims at aggressive activity. In order for a drive to be expressed, one needs an object. The drive itself is indifferent to the object; the drive simply seeks expression. So the aggressive drive will seek to vent itself on whatever object happens to be present.

We have already seen that drives do not just blindly impel an agent to act. Rather, drives operate by influencing the agent's perception and reflective thought, so that the agent sees a certain activity as warranted. With this in mind, suppose a drive is active, and seeks expression. If an appropriate object is unavailable, the drive will seek expression on whatever object happens to be present. The aggressive drive would most naturally be expressed upon things worthy of aggression. But, if there are no such objects, the drive will lead the agent to *seek* objects. So, a neutral stimulus may be interpreted as worthy of aggression, as *warranting* aggression. For example, the cashier's distraction may be seen as a personal snub, worthy of a rude remark. The driver's pulling in front of the car may be seen as an aggressive attack, worthy of horn play

and rage. I take it that this is a familiar phenomenon: anyone who has been in the grip of rage, jealousy, or any other strong affect can understand the sense in which these affects *seek* objects.

In this way, drives affect the agent's perceptions of reasons. The aggressive drive does not just produce a blind urge that causes the agent to act aggressively. Rather, the aggressive drive manifests itself by producing desires, affects, and perceptual saliences that jointly incline the agent to see aggression as warranted by the circumstances. This is why Nietzsche writes that a drive will "emphasize certain features and lines in what is foreign, in every piece of the 'external world', retouching and falsifying the whole to suit itself" (BGE 230). ↵

We can now put some points together. In the preceding section, we saw that drives manifest themselves by generating evaluative orientations. In this section, we have seen that when a drive is active it will induce a particular *kind* of orientation; it will induce an orientation that inclines the agent to take steps toward fulfilling the drive, by making it appear as if taking these steps is *warranted* by the situation at hand. For example, when the aggressive drive seeks to discharge itself, it will generate evaluative orientations that lead the agent to see aggressiveness as warranted by the situation at hand. So a drive manifests itself by impacting the agent's rational capacities.

And now we can begin to see something interesting: being moved by a drive and being moved by reflective thought are not distinct processes. Drives move us by directing and influencing our reflective thought.³² The [next section](#) examines this point in detail.

4 Drives and Reflective Agency

This essay began by posing three questions about drives. First, we asked what a drive is. We have seen that drives are a particular type of disposition, which manifests itself by generating an evaluative orientation. Second, we asked what type of awareness an agent has when she is being moved by a drive. We have seen that the agent typically lacks awareness of the end at which the drives dispose her to aim. With these claims in place, we can now turn to the third question: how should this understanding of drives impact our conception of reflective agency?

Nietzsche is notoriously skeptical of reflective agency. He explains paradigmatically reflective phenomena, such as self-conscious episodes of choice, as precipitates of drives. He tells us that reflection does not enable one to escape the influence of drives, arguing that an agent who acts reflectively is still "secretly guided and channeled" by his drives (BGE 3). In addition, he claims that whenever an agent steps back from and reflects upon a drive, the agent's "intellect is only the blind instrument of *another drive*" (D 109). These passages certainly seem designed to call into question our ordinary understanding of agency. In the following sections, I will ask how these passages should be understood.

4.1 Recent Interpretations of Nietzsche on Choice

We need an explanation of what Nietzsche means when he tells us that choice and reflective thought are guided or channeled by drives. Moreover, this explanation should make it clear why Nietzsche thinks the claim has implications for our understanding of the role of choice and reflective thought in agency. ↵

There seems to be a consensus within the Nietzsche literature on two points about choice. First, Nietzsche clearly denies that choice is a necessary condition for action. As we saw in section 1.3, Nietzsche claims that many actions occur without the agent's engaging in an episode of choice, indeed without the agent's even attending to her action. Second, it is by now a commonplace that Nietzsche rejects the *libertarian* conception

of choice, according to which an agent's choices are undetermined by prior events. As Gemes puts it, Nietzsche rejects "the notion of a will autonomous from the causal order, an uncaused cause" (2006: 325).

Of course, this leaves open a vast range of possible views about the nature of choice: few philosophers have claimed that choice is a necessary condition for action, and non-libertarian accounts of choice are legion. So these two claims about choice do not make Nietzsche's account seem particularly original.

A third claim, which has been defended by Brian Leiter, would make Nietzsche's account quite revolutionary. According to Leiter, Nietzsche argues that choice is epiphenomenal: "there is no causal link between the experience of willing and the resulting action" (2007: 13). On this interpretation, the agent's drives and other non-conscious factors³³ cause action, while the agent's reflective choices are simply idle.

Although Leiter's epiphenomenalist interpretation of choice would have dramatic implications, several commentators have argued on both textual and philosophical grounds that Nietzsche is not an epiphenomenalist.³⁴ I lack the space to examine these arguments here, but I do wish to register my agreement with them and to make one further point: Leiter's interpretation has textual costs. There are a number of passages in which Nietzsche appears to rely on the idea that choice can be causally efficacious, and Leiter's interpretation forces us to explain away these passages. For example, Nietzsche praises the "sovereign" or "autonomous" individual, who is distinguished by the fact that he "has his own independent, protracted will" (*GM* II: 2). Elsewhere, Nietzsche develops these ideas, claiming that "strong" agents have the power "not to react at once to a stimulus, but to gain control of all the inhibiting, excluding instincts...the essential feature is precisely not to 'will', to be able to suspend decision. All unspirituality, all vulgar commonness, depend on an inability to resist a stimulus: one must react, one follows every impulse" (*TI*: "What the Germans Lack" 6). In the same work, Nietzsche defines weakness as the "inability *not* to respond to a stimulus" (*TI*: "Morality as Anti-Nature" 2). The weak individual's actions are determined by whatever impulse or stimulus happens to arise; he possesses no capacity to direct his own behavior. By contrast, the strong individual is able to check his impulses and resist stimuli.³⁵ ↵

p. 750

In these passages, Nietzsche claims that some individuals have the capacity to control their behavior. Leiter's epiphenomenalist interpretation must treat these passages as rhetorical excesses or clumsy phrasings.³⁶ This is certainly possible: Nietzsche may have inadvertently invoked images of a causally efficacious capacity for choice, or his texts may be inconsistent. However, it would be preferable to find an interpretation of Nietzsche's views on choice that does not require us to discount any published passages. In the next section, I will show that such an interpretation is available.³⁷

4.2 Choice and the Possibility of Reflective Detachment

I am going to argue that Nietzsche is not primarily interested in questioning the causal connection between *choice* and *action*. Rather, Nietzsche's drive psychology problematizes the connection between *the agent* and *choice*.

To begin, let's consider the model of reflective choice that Nietzsche is attacking. I take it that Nietzsche is attacking a very influential model of agency, which is associated with Locke and Kant. The central claim of this model is that self-conscious reflection enables a deliberative suspension of motives. 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" (1975: 263). Kant endorses a similar model of deliberation, writing that human choice "can indeed be *affected* but not *determined* by impulses...*Freedom* of choice is this independence from being determined by sensible impulses" (*Metaphysics of Morals* 6: 213–14; see Kant 1996). Christine Korsgaard describes the Kantian model of deliberation as follows:

Our capacity to turn our attention to our own mental activities is also a capacity to distance ourselves from them, to call them into question...I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is this desire really a reason to act? (Korsgaard 1996: 93)

p. 751 The Lockean/Kantian claim about deliberative suspension can be broken into two parts. First, there is a claim about motivation: self-conscious reflection enables us to distance ourselves from our motives, thereby making these motives cease to “dominate” us or “suspending” these motives.³⁸ Second, there is a claim about normativity: once we have suspended a motive, we “consider” the motive, “weigh” it, ask if it is “really a reason to act.” In other words, once I have begun reflecting on the motive, what I do is look for a reason to act on the motive.

Of course, Locke and Kant do think that choice is causally efficacious; they believe that choice results in action. But they view choice as philosophically significant not simply because it results in action, but because in choosing, the agent suspends and rationally assesses her motives. So, one way of attacking this model would be to argue that choice doesn't cause action. But another way, which I think Nietzsche pursues, is to grant that choice causes action and attack the claim that in choosing, the agent suspends his motives.

For an illustration of the way in which Nietzsche is attacking this point, recall Murdoch's example, discussed in section 2.2. This example involves a mother, M, and a daughter-in-law, D. M initially sees D's behavior as brusque, rude, pert, and juvenile; later, she sees the same behavior as spontaneous, simple, gay, and delightfully youthful. A slight modification of Murdoch's example can be used to illustrate the complex way in which attempts at reflective detachment can fail. Let us imagine 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, at this earlier stage, M reflects on her dislike of D. Reasons for this dislike are forthcoming: D is vulgar and brusque. So reflection on the dislike apparently vindicates the dislike. But notice that the perception of D's behavior as vulgar and brusque is, in part, a result of M's jealousy. So the apparent detachment from her dislike of D, the reflective scrutiny of that dislike, the assessment of the dislike in light of evidence from observation, is influenced by another aspect of the very attitude from which M is attempting to detach.³⁹

p. 752 M is acting reflectively, and her choice does result in action. Yet M is in the thrall of attitudes that operate in the background. I suggest that this is what Nietzsche has in mind when he claims that reflection does not enable one to escape the influence of drives. The reflective agent is, in one sense, different from the unreflective agent: after all, the reflective agent deliberates, thinks about reasons for acting, and examines her motives, whereas the unreflective agent does none of this. But in another sense, the reflective agent is not so different from the unreflective agent: while the reflective agent supposes that she is escaping the influence of her drives, she is mistaken. The influence of the drive has simply become more covert.⁴⁰ ↵

If this is correct, then we needn't interpret Nietzsche as claiming that conscious choice is epiphenomenal. As mundane events such as choosing what to eat for breakfast indicate, our choices result in action all the time. But the causal connection between choice and action is not sufficient to demonstrate that we self-consciously control our actions, for self-conscious choice and the reflective assessment of motives are determined by drives that operate in the background. Reflectively choosing to A, and being caused to A by one's drives, are not distinct phenomena. *This* is the problem that Nietzsche's drive psychology raises: choice may control action, but agents do not control choice.⁴¹

5 Conclusion

A drive is a disposition that induces an evaluative orientation. Drives manifest themselves by structuring the agent's perceptions, affects, and reflective thought. Moreover, drives do not simply arise in response to external stimuli; they actively seek opportunities for expression, sometimes distorting the agent's perception of the environment in order to incline the agent to act in ways that give the drives expression.

This account of drives requires us to rethink our notion of deliberative or reflective agency. An agent who deliberates seems to enjoy a certain detachment from her motivational states. The deliberating agent experiences herself as capable of suspending the effects of her motivational states, and determining her action by choice. The drive psychology complicates this account. While it may be true that the agent who deliberates is not immediately compelled to act by her motivational states, her drives and other motives do continue to operate, in a subterranean fashion, even as the agent reflects on them. In many cases, the drives appear to decisively guide the agent's reflective choice in ways that she does not recognize.

This raises a potential problem. If the deliberating agent's thoughts and actions are guided, sometimes decisively, by her drives, can the actions that issue from her genuinely be regarded as *her* doings? Nietzsche sometimes suggests not: ^{p. 753} ↵

'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)

In this quotation, Nietzsche argues that many appearances of agency are illusory: what looks like a case of the agent's activity is better described as a case of the agent's being acted upon, presumably by her own drives. Echoing this point in another disconcerting passage, he writes, "Nothing is rarer than a *personal* action. A class, a rank, a race, an environment, an accident—everything expresses itself sooner in a work or deed, than a 'person'" (*WP* 866).

However, Nietzsche does not appear to believe that *every* action has this structure. After all, Nietzsche repeatedly speaks of "self-determination," "taking responsibility for oneself," and being a "sovereign individual" (*HAH* I: Preface 3; *TI*: "Skirmishes of an Untimely Man" 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.⁴²

Thus, there is some sense in which the agent acting under the influence of drives may be a passive conduit for the drives; however, Nietzsche also suggests that there is some way of acting that avoids this problem. Although I lack the space to address these topics here, my hope is that the analysis of drive psychology offered in this essay will bring us closer to understanding them.⁴³

Bibliography

(A) Works by Nietzsche

Reference edition of Nietzsche's works:

KSA Sämtliche Werke: Kritische Studienausgabe in 15 Einzelbänden, ed. G. Colli and M. Montinari. Berlin: De Gruyter, 1988.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

List of abbreviations of Nietzsche's works:

A The Antichrist, trans. W. Kaufmann. New York: Viking, 1954.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

BGE Beyond Good and Evil, trans. W. Kaufmann. New York: Modern Library, 1968.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

CW The Case of Wagner, trans. W. Kaufmann. New York: Random House, 1967.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

D Daybreak, trans. R. J. Hollingdale. Cambridge: Cambridge University Press, 1982.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

EH Ecce Homo, trans. W. Kaufmann. New York: Modern Library, 1968.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

GM On the Genealogy of Morality, trans. W. Kaufmann. New York: Modern Library, 1968.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

GS The Gay Science, trans. W. Kaufmann. New York: Vintage, 1974.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

p. 754 *HC "Homer's Contest,"* trans. W. Kaufmann. New York: Viking, 1954. ↪

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

HAH Human, All Too Human, trans. R. J. Hollingdale. Cambridge: Cambridge University Press, 1986.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

NCW Nietzsche Contra Wagner, trans. W. Kaufmann. New York: Viking, 1954.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

TI Twilight of the Idols, trans. W. Kaufmann. New York: Viking, 1954.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

UM Untimely Meditations, trans. R. J. Hollingdale. New York: Cambridge, 1997.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

WP The Will to Power, trans. W. Kaufmann and R. J. Hollingdale. New York: Vintage, 1967.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Z Thus Spoke Zarathustra, trans. W. Kaufmann. New York: Viking, 1954.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

(B) Other Primary Works

WWR Schopenhauer, Arthur. 1969. *The World as Will and Representation*, trans. E. E. F. J. Payne (2 vols). New York: Dover.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

(C) Other Works Cited

Anscombe, G. E. M. 2000. *Intention*. Cambridge, Mass.: Harvard University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Assoun, Paul-Laurent. 2000. *Freud and Nietzsche*, trans. R. L. Collier. New Brunswick: Athlone Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Boring, Edwin G. 1929. *A History of Experimental Psychology*. New York: D. Appleton-Century Company.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Brewer, Talbot. 2002. "Maxims and Virtues," *Philosophical Review* 111: 539–72.

[Google Scholar](#) [WorldCat](#)

Clark, Maudemarie. 1998. "On Knowledge, Truth, and Value: Nietzsche's Debt to Schopenhauer and the Development of his Empiricism," in C. Janaway (ed.), *Willing and Nothingness: Schopenhauer as Nietzsche's Educator*. Oxford: Oxford University Press, 37–78.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Clark, Maudemarie and Dudrick, David. 2009. "Nietzsche on the Will: An Analysis of BGE 19," in Ken Gemes and Simon May (eds), *Nietzsche on Freedom and Autonomy*. Oxford: Oxford University Press, 247–68.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Conway, Daniel. 1997. *Nietzsche's Dangerous Game: Philosophy in the "Twilight of the Idols."* Cambridge: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Cox, Christoph. 1999. *Nietzsche: Naturalism and Interpretation*. Berkeley and Los Angeles: University of California Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Darwin, Charles. 1993. *On the Origin of Species*. New York: Modern Library.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Darwin, Charles. 2004. *The Descent of Man*. London: Penguin.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Freud, Sigmund. 1957. *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, ed. J. Strachey. London: Hogarth Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Gemes, Ken. 2001. "Post-Modernism's Use and Abuse of Nietzsche," *Philosophy and Phenomenological Research* 52: 337–60.

[Google Scholar](#) [WorldCat](#)

Gemes, Ken. 2006. "Nietzsche on Free Will, Autonomy and the Sovereign Individual," *Proceedings of the Aristotelian Society* 80 (supp. vol.): 321–38.

[Google Scholar](#) [WorldCat](#)

Hales, Steven and Welshon, Rex. 2000. *Nietzsche's Perspectivism*. Champaign, Ill.: University of Illinois Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Huxley, T. H. 1874. "On the Hypothesis that Animals are Automata, and its History," *Fortnightly Review* 95: 555–80.

[Google Scholar](#) [WorldCat](#)

James, William. 1890. *The Principles of Psychology*. New York: Dover.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Janaway, Christopher. 2006. "Nietzsche on Free Will, Autonomy and the Sovereign Individual," *Proceedings of the Aristotelian Society* 80 (supp. vol.): 339–57.

[Google Scholar](#) [WorldCat](#)

p. 755 Janaway, Christopher. 2007. *Beyond Selflessness: Reading Nietzsche's Genealogy*. New York: Oxford University Press. ↵

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Kant, Immanuel. 1996. *The Metaphysics of Morals*, ed. Mary Gregor. New York: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*, ed. Mary Gregor. New York: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Katsafanas, Paul. 2005. "Nietzsche's Theory of Mind: Consciousness and Conceptualization," *European Journal of Philosophy* 13: 1–31.

[Google Scholar](#) [WorldCat](#)

Katsafanas, Paul. 2011. "The Concept of Unified Agency in Nietzsche, Plato, and Schiller," *Journal of the History of Philosophy* 49.1: 87–113.

[Google Scholar](#) [WorldCat](#)

Katsafanas, Paul. 2012. "Nietzsche on Agency and Self-Ignorance," *Journal of Nietzsche Studies* 43.1: 5–17.

[Google Scholar](#) [WorldCat](#)

Korsgaard, Christine. 1996. *The Sources of Normativity*. New York: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Leiter, Brian. 2002. *Nietzsche on Morality*. London: Routledge.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Leiter, Brian. 2007. "Nietzsche's Theory of the Will," *Philosophers Imprint* 7.7: 1–15.

[Google Scholar](#) [WorldCat](#)

Leiter, Brian. 2011. "Who is the Sovereign Individual? Nietzsche on Freedom," in Simon May (ed.), *Nietzsche's "On the Genealogy of Morality": A Critical Guide*. Cambridge: Cambridge University Press, 101–19.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Locke, John. 1975. *An Essay Concerning Human Understanding*. New York: Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Lowry, Richard. 1971. *The Evolution of Psychological Theory*. Hawthorne, NY: Aldine Publishing.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Moore, Gregory. 2002. *Nietzsche, Biology, and Metaphor*. New York: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Murdoch, Iris. 1985. *The Sovereignty of Good*. New York: Routledge.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Parkes, Graham. 1996. *Composing the Soul*. Chicago: University of Chicago Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Poellner, Peter. 1995. *Nietzsche and Metaphysics*. New York: Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Proudhon, Pierre-Joseph. 1994. *What is Property?* New York: Cambridge University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Richards, Robert. 1987. *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: University of Chicago Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Richardson, John. 1996. *Nietzsche's System*. New York: Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Richardson, John. 2004. *Nietzsche's New Darwinism*. New York: Oxford University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Rosenfield, Lenora. 1968. *From Beast-Machine to Man-Machine: Animal Soul in French Letters from Descartes to La Mettrie*. New York: Columbia University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Schacht, Richard. 1983. *Nietzsche*. New York: Routledge.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Schopenhauer, Arthur. 1969. *The World as Will and Representation*, trans. E. F. J. Payne. New York: Dover.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Stampe, Dennis. 1987. "The Authority of Desire," *Philosophical Review* 96: 335–81.

[Google Scholar](#) [WorldCat](#)

Thiele, Leslie Paul. 1990. *Friedrich Nietzsche and the Politics of the Soul*. Princeton: Princeton University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Thorpe, W. H. 1956. *Learning and Instinct in Animals*. London: Methuen.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Wilm, E. C. 1925. *The Theories of Instinct*. New Haven: Yale University Press.

[Google Scholar](#) [Google Preview](#) [WorldCat](#) [COPAC](#)

Notes

- 1 Nietzsche seems to regard *Instinkt* and *Trieb* 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 *Instinkt* and *Trieb*. (I use *drive* instead of *instinct* because, we will soon see, the English term "instinct" has misleading connotations.) Daniel Conway claims that Nietzsche distinguishes *Instinkt* and *Trieb* beginning in his works of 1888. According to Conway, beginning in *Twilight*, *Instinkt* refers to a *Trieb* that has been "organized" or "trained to discharge" in a specific way (1997: 30–4). I find Conway's textual evidence for this alleged distinction unpersuasive; however, we need not resolve the issue here, for this distinction would not affect the points that I make in the text.

- 2 In particular, Nietzsche contrasts his drive psychology with the accounts of “clumsy naturalists who can hardly touch on ‘the soul’ without immediately losing it” (*BGE* 12).
- 3 For some examples, see *KSA* 12: 1 [58]; *WP* 481, 260, 567.
- 4 “That a psychologist without equal speaks from my writings, is perhaps the first insight reached by a good reader” (*EH*: “Why I Write Such Good Books” 5).
- 5 While Clark and Dudrick provide some intriguing remarks about drives in this article, their primary focus is elsewhere: they attempt an analysis of the notion of the will. For this reason, it is not entirely clear whether the remarks on drives constitute their considered view.
- 6 Though Poellner interprets drives as agents, he is quite sensitive to this problem. He notes that certain passages in Nietzsche’s work suggest that drives should *not* be understood “as themselves conscious of their activity—of their desiring, interpreting, willing, commanding, and obeying” (1995: 216). Accordingly, he seeks an explanation of how drives could be non-conscious agents. Failing to find a satisfactory explanation, he concludes, with admirable candor, that Nietzsche’s remarks on unconscious drives are ultimately indefensible (1995: 215–29). Clark and Dudrick also address this problem; see the next note.
- 7 Arguably, there could be some advantage in explaining agency in terms of simpler, less complex sub-agents. Thus, if we could understand drives as less complex agents, perhaps the drive psychology would have some explanatory power. Clark and Dudrick (2009) give an interesting argument in favor of this point. They claim that if we interpret drives as simple sub-agents, capable of commanding and obeying one another, then we can make sense of more complex agential phenomena such as valuing and resisting temptation. While this proposal is intriguing, I think there are grounds for objection. First, it is not clear that commanding and obeying are simpler activities than valuing and resisting temptation. Moreover, commanding and obeying require, at the very least, the presence of consciousness. So the type of agency attributed to drives is still quite robust. Thus, while Clark and Dudrick’s proposal is certainly an improvement upon the other homuncular views, it still faces certain problems.
- 8 Richardson addresses the same question in a later work, asking whether Nietzsche has “a viable notion of drives...At issue, in particular, will be how Nietzsche can attribute the *end-directed* character he clearly does to these drives and wills, without illicitly anthropomorphizing an implausible mentality into them.” (2004: 13)
- 9 I say a “typical” ascetic because natural selection may not have disposed *every* agent toward sexual activity.
- 10 Richardson does anticipate this form of objection, and responds as follows:

We should bear in mind that this valuing need not—and principally does not—occur in a conscious act...We suppose that “our values” are those we put into *language* and *consciousness*...But according to Nietzsche...the really effective or influential values are not those conscious ones...Values are built into our bodies, and their conscious and linguistic expression is something quite secondary. (2004: 73–4)

Part of Richardson’s point, here, is that Nietzsche might reject my characterization of the ascetic, above. In particular, Nietzsche might deny that the ascetic’s conscious thoughts about the disvalue of sex and the value of celibacy actually count as *values*. Nietzsche might, instead, take the fact that the ascetic is disposed toward sex to indicate that he values it. I agree with Richardson that this is one way of interpreting the texts: we could read Nietzsche as departing from our ordinary conception of value, and introducing this novel conception of value, according to which values are identified with selected dispositions. However, I think there is some reason to resist this interpretation. Arguably, it is an essential feature of our concept of value that values can conflict with motivations in general and dispositions in particular. In the following sections, I will argue that there is another way of interpreting Nietzsche’s remarks about drives’ valuing, which preserves a distinction between being disposed to E and valuing E.

- 11 Many contemporary philosophers accept Anscombe’s thesis that if an agent intentionally *As*, then the agent knows that she is *A-ing* (Anscombe 2000). But this should not be mistaken for the claim that if an agent intentionally *A-s*, then the agent’s *A-ing* is an object of explicit attention. Just as I can know straightforward factual matters, such as my birth date, without explicitly attending to them at all times, so too I can know that I am walking to my office without attending to the movements of my legs.
- 12 In addition, the idea that we lack epistemic access to our causally effective motives (at least in ordinary circumstances) is a frequently voiced theme in Augustine, La Rochefoucauld, Montaigne, Spinoza, and Schopenhauer, to name but a few.
- 13 Some principal figures in the debate were Le Roy, Reimarus, Condillac, Erasmus Darwin, Cabanis, Cuvier, Flourens, Lamarck, and Charles Darwin. On the more philosophical side, instinct was discussed by Schiller, Hegel, and Schopenhauer. Turning to literature, the most obvious influences on Nietzsche are Hölderlin and Emerson, both of whom frequently employed the notion of instinct. For useful discussions of the history of the instinct concept, see Richards 1987, Thorpe 1956, and Wilm 1925. Boring 1929 and Lowry 1971 provide more general discussions of the history of psychology. In regard to Nietzsche, Moore 2002, Parkes 1996 and Assoun 2000 are particularly helpful.

- 14 The claim that learned behaviors require awareness of the action's goal is perhaps most plausible when we are considering isolated animals that engage in highly original behaviors. Otherwise, we can imagine one animal simply copying or imitating the behavior of another, in much the way that an infant might imitate the gestures of its parents without understanding their purpose. See also note 17.
- 15 See, for example, Fontaine's account of the Port Royal experimenters: "they said the animals were clocks; that the cries they emitted when struck were only the noise of a little spring that had been touched, but that the whole body was without feeling" (*Mémoires pour servir à l'histoire de Port-Royal*, 1738, quoted in Rosenfield 1968: 54).
- 16 In place of "is not aware," we might substitute "does not know," "does not believe," "is not cognizant of the fact," and so on. I intend (1) to be neutral among these formulations.
- 17 Brian Leiter and Ken Gemes (personal communication) point out that some philosophers may find these characterizations of animal awareness problematic. Why is the cuckoo's excitement at the perception of a nest a *motive*, rather than a mere *cause*? Why say that the cuckoo has thoughts involving *nests*—and indeed, can it even have the concept of a nest? While these questions are important, for present purposes we can set them aside. I am here discussing the way in which certain eighteenth- and nineteenth-century thinkers characterized animal thought and motivation. Some thinkers—those embracing claim (1)—would deny that the cuckoo has motives, is aware of nests, and so forth. Other thinkers—those embracing claim (2)—wish to describe these animal actions in terms of more advanced mental processes. My task at this stage of the essay is not to *assess* these competing claims, but simply to *distinguish* them.
- 18 To clarify what proponents of (2) have in mind, it is worth noting that there are relatively straightforward ways of testing whether an animal knows that it is *A*-ing in order to *G*: we can break the connection between *A*-ing and *G*-ing, and see whether the animal continues to *A*. The rat that learns to press a lever a certain number of times in order to acquire a food pellet will cease to do so, if the lever-pressing fails to yield food consistently. The wasp which collects grubs and stores them beside its eggs will *not* cease to do so if, for example, its eggs are clearly destroyed. So the rat seems to know that it is pressing the lever in order to obtain food, whereas the wasp seems not to know that it is collecting grubs in order to feed its offspring.
- 19 To be precise, Schopenhauer does seem to allow that in some cases, drives move a person by blindly impelling him to act. Schopenhauer's claim that certain instinctive actions bear "a remarkable similarity to those of somnambulists" (*WWR* II: 344) seems to make this point. However, when Schopenhauer offers extended discussions of the operations of drives, he typically treats them as operating *through*, rather than *independently of*, the agent's reflective thoughts.
- 20 There are, however, two potential problems with this claim. First, it is not clear what criteria Schopenhauer employs in order to determine that reproduction is the true purpose of the action, rather than, say, romantic love. To be sure, reproduction is what the action was *selected* for; but it is not obvious why the selected purpose of the instinctual activity should be identified with the true purpose. Second, it is worth noting that the agent is not straightforwardly *wrong* about his actions: it may well be true that his love will provide him with immense happiness. His descriptions of his own actions need not be false, but they are *incomplete*. This raises a question: what if the agent simply doesn't care that his knowledge is incomplete? After all, knowledge is always incomplete; why should knowledge of one's own actions be any different? In short, why should Schopenhauer's remarks trouble us? For the moment, I want to bypass this question; what concerns us here is not the particular purpose that Schopenhauer singles out, but rather the *structure* of the action that Schopenhauer discusses. That is, what concerns us is not the claim that an agent is actually pursuing reproduction, but the more general claim that an agent takes himself to be pursuing *A*, whereas "pursuing *B*" is a more apt description.
- 21 For a related discussion, see Clark 1998: §3.
- 22 Schopenhauer also discusses this idea throughout chapter XIX of *WWR* II. For contemporary discussions of related ideas, see for example Stampe 1987 and Brewer 2002.
- 23 Nietzsche must have regarded this point as extremely important, for he discusses it in nearly identical fashion in several of his major works. For example, *Twilight* contains a virtually identical passage, in which nothing but the example has changed (*TI*: "The Four Great Errors" 4).
- 24 For a discussion of Nietzsche's remarks on conceptualized experiences, see Katsafanas 2005.
- 25 The role that Murdoch gives to self-scrutiny and self-criticism is roughly analogous to the role that Nietzsche gives to self-understanding and genealogy: in both cases, the agent discovers the impact of hidden or unnoticed motives, and thereby puts herself in a position to counteract them. Murdoch's account is, however, somewhat simpler and more straightforward than Nietzsche's. For example, notice that Murdoch simply assumes that *M*'s latter view is in some way *better* than the earlier one—that the latter view is *correct* in a way that the former view is not. As Nietzsche would point out, there is no reason to take this as obvious. The earlier view was motivated by jealousy; the latter view might be motivated by a desire to please her son. The mere fact that a view changes in such a way that it presents its object in a more appealing light does not imply that the new view is more adequate; the reverse could be the case. While Nietzsche does think that there are better and worse views, he recognizes that the question of whether a given view is better or worse than others is always difficult, and cannot always be determined.

26 Thus, embodied drives can reason/evaluate/interpret in the sense that they can induce in agents affective dispositions that constitute reasonings/evaluations/interpretations.

27 Notice that even non-self-conscious animals could be said to reason, evaluate, and interpret, in the above sense. That is, if Nietzsche's talk of reasoning, evaluating, and interpreting is intended to express the fact that many animals have affective orientations, then it makes perfectly good sense.

28 I have discussed this point in Katsafanas 2012.

29 It is important to note that an organism need not and typically will not be aware of the evaluative outlook manifested in its orientation toward its environment. A self-conscious animal, such as a human being, *can* become aware of the partiality and selectivity in its orientation, as Murdoch's example demonstrates. But this takes work. Typically, agents will be largely ignorant of their own evaluative outlooks.

30 Ken Gemes defends a related claim about drives in his contribution to this volume. See especially section 2.1. There, Gemes argues that Nietzsche's claims about drives "interpreting" phenomena and generating "perspectives" are best understood as claims about drives leading agents to interact with their environments in determinate ways.

31 The Standard Edition translates the title *Triebe und Triebchicksale* as "Instincts and their Vicissitudes." I think *Trieb* is better translated as *drive* in this context. In the quotations from Freud above, I follow the Standard Edition, but translate all occurrences of *Trieb* as *drive*.

32 Although I lack the space to argue for this point here, Nietzsche seems to embrace the following claims: (1) drives *always* exert some influence upon reflective thought, but (2) the *extent* of this influence differs from case to case, and, perhaps most importantly, (3) the drives' influence *often*, but *not always*, undermines the agent's claim to being in control of the action. I return to the third point, briefly, in the final paragraphs of this essay. For an extended defense of the third point, see Katsafanas 2011.

33 In particular, Leiter claims that actions are caused by "type-facts" about the person, where "type-facts, for Nietzsche, are either *physiological* facts about the person or facts about the person's unconscious drives or affects" (2007: 7).

34 See, for example, Clark and Dudrick 2009, Gemes 2006, and Janaway 2006. I discuss this issue in Katsafanas 2005.

35 There are a number of similar passages, both in the notebooks and in the published works. In *WP* 95, Nietzsche condemns nineteenth-century thinkers for being "deeply convinced of the rule of cravings. (Schopenhauer spoke of 'will'; but nothing is more characteristic of his philosophy than the absence of all genuine willing)." *WP* 928 speaks of great individuals controlling their affects: "Greatness of character does not consist in not possessing these affects—on the contrary, one possesses them to the highest degree—but in having them under control." *WP* 933 makes a similar point: "*In summa: domination of the passions, not their weakening or extirpation!*—The greater the dominating power of a will, the more freedom may the passions be allowed. The 'great man' is great owing to the free play and scope of his desires and to the yet greater power than knows how to press these magnificent monsters into service." *WP* 962 claims that a great individual "has the ability to extend his will across great stretches of his life." For an example from a published work, consider *Zarathustra*. In *Z II*: "Of Redemption" and elsewhere, Zarathustra claims that the *past* is a source of dissatisfaction because it cannot be modified by the will: "powerless against what has been, the will is an angry spectator of all that is past. The will cannot will backwards...that is the will's loneliest melancholy." If the will could not will *forward*, either—if acts of will had no causal impact on the agent's actions—then this claim would be unintelligible. The future would be just as inaccessible as the past.

36 And indeed, Leiter attempts to deflate these passages in Leiter 2011.

37 To be clear, we should distinguish two claims: (1) Nietzsche's account of the relationship between drives and conscious thought is compatible with the claim that conscious thought is causally efficacious; (2) Nietzsche maintains that conscious thought is causally efficacious. I think both (1) and (2) are true—this, I submit, is the most natural reading of the passages cited above. However, in this essay I argue only for claim (1). For defenses of claim (2), see Katsafanas 2005 and the works cited in note 34.

38 To be precise, Kant and Locke seem to maintain that for any motive, I can self-consciously suspend its influence. We should distinguish this claim from the much stronger claim that I can, at any moment, suspend the influence of *all* of my motives. (Compare the analogous point about belief: we can, at any moment, self-consciously reflect on and critically assess any given belief; but we cannot, at one moment, self-consciously reflect on and critically assess *all* of our beliefs. For we assess given beliefs in light of other beliefs, which are held fixed.)

39 Of course, in Murdoch's original example M eventually does achieve a reflective detachment from her jealous motives. However, achieving this reflective detachment requires hard, self-critical work: M needs to engage in an extended investigation of the effects of various attitudes on her perceptions.

40 Although I lack the space to argue for this point here, I think Nietzsche maintains that reflection and self-understanding enable an agent to counteract the effects of particular drives. Murdoch's original example illustrates this point: by coming to understand the way in which her jealousy influences her reflection, M manages to counteract its effects. Indeed, this seems to be one reason why Nietzsche emphasizes the importance of self-understanding: self-understanding enables one

to counteract the effects of certain drives, and thereby renders the agent increasingly in control of her action. See note 32. For a related discussion of the role of self-understanding, see Richardson 2004 (especially 95–103).

- 41 Section 4.1 noted that Leiter interprets Nietzsche as an epiphenomenalist about willing. However, Leiter acknowledges that there is some evidence for a different reading of Nietzsche, according to which “the will is, indeed, causal, but it is not the *ultimate* cause of an action: something *causes* the experience of willing and then the will causes the action” (Leiter 2007: 13). Leiter calls this the “Will as Secondary Cause” interpretation. The interpretation that I have defended, above, is in some respects similar to Leiter’s Will as Secondary Cause reading. However, there are two potential disagreements that merit attention. First, the Will as Secondary Cause reading suggests a unidirectional causal path between drives and choice: drives determine conscious choices, which then determine actions. On my interpretation, however, there is a bidirectional causal path: drives causally influence conscious thoughts, but conscious thoughts (including choices) also causally influence drives (see Katsafanas 2005 for the details). Second, we should distinguish two readings of the relationship between drives and conscious thoughts. According to the strong reading, drives *determine* conscious thoughts: if two agents’ drives (and circumstances) are identical, then their conscious thoughts must be identical. According to the weak reading, drives merely *influence* conscious thought: two agents’ drives (and circumstances) could be identical, and yet they could have different conscious thoughts. I accept the weak reading, whereas Leiter’s Will as Secondary Cause interpretation seems to be premised upon the strong reading. For a defense of these points, see Katsafanas 2005 and 2011.
- 42 Katsafanas 2011 provides an argument for the claim that Nietzsche has a conception of genuine agency.
- 43 For helpful discussions of the material in this essay, I owe great thanks to Lanier Anderson, Ken Gemes, Christine Korsgaard, Brian Leiter, Richard Moran, Bernard Reginster, John Richardson, Mathias Risse, and Danielle Slevens.