Against Nietzsche's 'Theory' of the Drives

ABSTRACT: *Nietzsche, we are often told, had an account of 'self' or 'mind' or a 'philosophical psychology', in which what he calls our 'drives' play a highly significant role. This underpins not merely his understanding of mind— in particular, of consciousness and action—but also his positive ethics, be they understood as authenticity, freedom, (self-)knowledge, autonomy, self-creation, or power. But Nietzsche did not have anything like a coherent account of 'the drives' according to which the self, the relationship between thought and action, or consciousness could be explained; consequently, he did not have a stable account of drives on which his positive ethics could rest. By this, I do not mean that his account is incomplete or that it is philosophically indefensible: both would leave open, misleadingly, the possibility of a rational reconstruction of Nietzsche's views; both would already assume more unity and coherence than we find in his texts. Specifically, as I show through detailed analysis, Nietzsche provides varied and inconsistent accounts of (1) what a 'drive' is, (2) how much we can know about drives, and (3) the relationship between drives and conscious deliberations about action. I conclude by questioning the hunt for a Nietzschean theory: is this the best way to be reading him?*

KEYWORDS: history of philosophy, nineteenth-century philosophy, German philosophy, philosophy of mind, action, consciousness, Nietzsche, drives

Nietzsche, it is often suggested, had an account of the 'self' or the 'mind' or a 'philosophical psychology' in which what he calls our 'drives' play a highly significant role. This underpins not merely his understanding of mind— in particular, of consciousness and action—but also his positive ethics, be they understood as authenticity, freedom, (self-)knowledge, autonomy, self-creation, or power.[1] But, as this paper argues, Nietzsche did not in fact have anything like a coherent account of 'the drives', according to which the self, the relationship between thought and action, or consciousness could be explained; consequently, he did not have an account of the drives on which his positive ethics could rest. By this, I do not mean that his account is incomplete or that it is philosophically indefensible or even that his writings move ambiguously between some two distinct

variants: all would leave open—wrongly, as I argue—the possibility of a rational reconstruction of Nietzsche's views; all would already assume more unity and coherence than we find in his texts. I make no objection to rational reconstruction in the history of philosophy or in Nietzsche scholarship. But not every writer's claims on every subject can be rationally reconstructed; here, such a reconstruction cannot be done. For the things Nietzsche says about drives—what they are, what we can know about them, and what their relationship to actions is—are so deeply and centrally conflicting in so many ways that no coherent position can be formed without ignoring significant published passages. What Nietzsche did have was a very clear picture of the target he wanted to attack. It will be helpful to begin there.

1. Nietzsche's Target: The 'Socratic' Picture

Nietzsche names Socrates and Plato as 'heirs' to the sort of 'primeval delusion' he confronts with his writings about mind (D 116[2]). In fact, one particular 'decision' that Socrates made bothered both Plato and Nietzsche: opting to die instead of escaping and running away (see *Crito, Phaedo*; see also TI, 'Socrates'). As Socrates awaits his death, he speaks of his disappointment with Anaxagoras, who had initially attracted Socrates by appearing to explain the world in terms of a ruling Mind. But, upon reading Anaxagoras's book, '[Socrates] found him making no use of Mind, not crediting it with any causality' (*Phaedo* 98b). Anaxagoras, Socrates suggests, might as well explain Socrates staying in prison by referring to his 'bones and sinews'. The true cause, Socrates insists, is his mind. The reason: 'I fancy these same sinews and bones would long since have been somewhere in Megara or Boeotia . . . if I had not thought it right and proper to submit to the penalty appointed by the State' (98e). Socrates does not deny that bones and sinews are necessary conditions for action: 'If I did not possess things like that—bones and sinews—I shouldn't be able to do what I had resolved upon'; it's just that 'to call things like that [i.e., necessary conditions] causes is quite absurd' (99a).

On Socrates's account of the relationship between what he thought and what he did, the following seems to be true: (1) Socrates could have chosen to do different things; (2) Socrates correctly weighed up his options, and (3) he acted according to which option he thought was right; (4) having acted, Socrates's motivations are plain for him to see and to be made available to others such that (5) he can be judged morally based on what he chose to do. Throughout this paper, I shall refer to this cluster of notions about the relationship between conscious deliberation and action as the 'Socratic picture'. ('Socratic' is a useful label. I do not say that Nietzsche always had Socrates in mind. Indeed, there are good reasons to think that Plato himself did not adhere to the 'Socratic picture', even in the early dialogues such as the *Crito,* see also Harte [2005].)

Anybody familiar with Nietzsche's remarks about drives (and related matters) knows that, in various places, he attacks each one. In as much as (1) suggests a 'free

---

2 References to Nietzsche's texts follow standard textual abbreviations. Translations are my own when stated; otherwise they refer to translations in the bibliography.

will' or mind as 'first cause', this would be a target of Nietzsche's scorn for more or less his whole career. Against (2), Nietzsche doubts that deliberators have reliable, epistemic access to the outcomes of their would-be choices (D 129). Against (3), he complicates the apparently simple causal link between deciding and doing (D 119, 120, 129, TI 'Errors', A 39). Against (4), he stresses both the unknowability of actions and the unreliability of conscious thought as a guide to ourselves (D 115, 116, 119, 120; GS 11, 335). For obvious reasons, Nietzsche frequently takes his various attacks on (1) through (4) to deal with (5), too (D 129, GS 335).

Whether these criticisms still speak to us will depend, first, on how widespread we think the 'Socratic picture' is in philosophy, psychology, folk psychology, and beyond. It will also depend on particular arguments that Nietzsche gives or on his acutely honest, novelistic descriptions of how decisions actually feel (see Leiter [2009] for analysis of some of his descriptions and corresponding arguments). His opposition to the Socratic picture remains throughout, and Nietzsche often combines an attack on part of it with a suggested alternative model. It *may* be that one or two of these suggested alternatives, taken out of context and given substance by other philosophers or psychologists, could form a plausible and challenging account. One aim of this paper is to emphasize how many different horses Nietzsche backs: it would hardly be surprising if one of their descendants went on, years later, to win a race (though I disagree that one of their descendants winning a race would have anything to do with claiming that this was what Nietzsche himself *meant* in some final sense—cf. Leiter 2009: 123). But as for Nietzsche himself: contrary to the scholarly consensus, there is no hope for a single, unified, or considered position. The very most we can get amounts to two coarse claims: first, that we act as a result of our 'drives', which battle it out between them to determine what we do; second, that our so-called 'conscious deliberations' do not accurately reflect this battle of the 'drives'. To be sure, this looks like the makings of a substantive philosophical position. But I suggest we would hesitate to accord it such status before we had clarification on three key points: (a) what sorts of things 'drives' are; (b) how, and how much, we know about them; and (c) what the relationship is between the battle of the drives, which really accounts for actions, and the conscious deliberations that, on the Socratic picture, seem to account for actions?

Nietzsche gives explicitly inconsistent answers regarding each of these three key points. Not that he fails to answer them: he *does* answer them, just in different, incompatible ways. Often—and this important point is easy to miss—he goes beyond merely stating inconsistent positions: he actually gives *arguments* in favor of the various contradictory positions. Unsurprisingly, then, plenty of his answers regarding the relationship between these key points are also inconsistent: we shall see, for example, that claims about unknowability of drives clash (i.e., regarding (b), above) with others about their functioning (i.e., regarding (c), above). This points to a complete failure to produce anything like an account we can work with, *even prior* to assessing it on philosophical grounds (and, as some have argued, the prospects for many proposed positions are poor—see Gardner [2009], Katsafanas [2013], and Anderson [2013]). The first two parts of this paper deal with the first two questions. The next three parts deal with the third question.

(Another exegetical option would be to construct a speculative Nietzschean psychology based on his other philosophical–especially normative–commitments. My focus is on what he did say, not on what he should say. But I have reasons for pessimism about the prospects of such an approach, generally, because a firm Nietzschean ground is hard to find and also for specific, independent reasons: his ethical commitments lead to even more complex psychological demands. For more on this see Stern (forthcoming).)

1. What is a Drive?

I have already implied more unity to Nietzsche's view than we find when we look at the texts. For, at first glance, the things that carry out the sort of role I have allotted to 'drives' may be called 'drives', 'instincts', 'affects', or 'tendencies'. These are prima facie very different things and (with the exception of 'drive' and 'instinct') would have been in the usage of Nietzsche's day. A closer look at these terms in general and then at their usage in Nietzsche will show, to a greater degree than has been acknowledged, the challenge that faces the seeker of a drive theory in Nietzsche.

Nietzsche seems to have taken the terms 'drive' and 'instinct' from the philosophical-psychological language of his day (see Katsafanas [2013] for selected cases; on his reading of Nietzsche and drives, see below). Despite the differences in their use in ordinary language, these terms may be treated as equivalent: for animals presumed to be without conscious deliberation, they appear to denote those forces that impel the animal to act in a way that *looks* rational and purposive but for which reason and purpose, in the animal itself, are evidently lacking. Hence, the bee constructs the hive because of its instincts: that is, it did not plan, consciously, to do so, but if the bee had consciously planned the hive, it couldn't have done a better job. The 'impulsion' is also important, since the drive pressures or propels the animal to seek out opportunities for the drive to express itself. Nietzsche knew such definitions. David Strauss, in the book Nietzsche reviewed for his first *Untimely Meditation* (UM I), cites Reimarus's account approvingly as follows: 'one observes in all animals certain drives, instincts or efforts' [*Triebe, Instinkte oder Bemühungen*] by means of which they can 'perform, masterfully, without any of their own consideration, experience or practice, without any instruction, example or model, that which the most perfect reason might have advised them to do for their own good' (Strauss 1895: 143, my translation; also 1895: 77). Where does the as-if reason come from? Pre-Darwinian explanations, as Strauss comments, simply but problematically transpose the 'subject' elsewhere—to God (Reimarus), the unconscious (Hartmann), or, he might have added, the will (Schopenhauer). Darwin, for Strauss, provides a better answer, since an instinct is an 'acquisition' from countless generations of natural selection: there is no hidden consciousness or subject to be found (Strauss 1895: 144–5). In the human case, drive or instinct may therefore be taken to be a nonconscious urge or guiding power that makes a person act in a way that *seems* rational and purposive but that is not in fact (consciously) rational and purposive.

A '*Hang*'—variously translated as a 'tendency', 'disposition', 'inclination', 'propensity', 'bent', or plain 'desire'—suggests, as these translations indicate, what someone or something tends to do. Nietzsche certainly appears to use *Hang* and 'drive' or *Hang* and instinct interchangeably in some cases. Often, he connects *Hang* with a 'taste' for something (BGE 2, 239, 245, GS 87). What someone happens to have a taste or *Hang* for—luxury or democracy or the unconditional (Nietzsche's examples)—doesn't look to be tied to biological needs, as might be suggest by 'drive' and 'instinct'.

The German term, '*Affekt*', suggests a forceful inner stirring or emotion—hence an '*Affekthandlung*' (an 'affect-action') is an action that is emotionally charged, rash, ill-considered. This is the sense that Kant harnesses when he distinguishes '*Affekt*' from '*Leidenschaft*' (passion): the former, not the latter, indicates a hasty, inner storm that soon passes, leaving one exhausted (Kant 1996: 536; 1997: 287–8). William James recognizes '*Affekt*' as a German philosophical term, which he translates as 'emotion', specifically suggesting a 'general seizure of excitement' (James 1894: 358). Unlike in the case of drive, instinct or '*Hang*', there is no reason to suppose that an affect endures in any meaningful sense. A hunger drive, one supposes, drives the animal toward the right place for food and is in some sense active throughout its life. And if you have a *Hang* for luxury or for democracy, then you are inclined toward them whenever they arise, other things being equal. But there is no reason to suppose that an 'affect' *endures*; indeed, for Kant, 'affects' are *characteristically* brief, which is why they belong to 'nonvirtue' rather than 'vice' (Kant 1997: 287). 'Tendency' and 'affect' are sufficiently distinct, for Kant, that he can speak happily of a 'propensity to an affect [*ein Hang zum Affekt*]', that is, a tendency to get a certain forceful stirring of inner emotion, like rage (Kant 1997: 287).

It might seem, now, as though we have a relatively clear way of distinguishing these terms: drive and instinct as biological, quasi-rational, and perhaps, on Darwinian grounds, inherited through generations of natural selection; 'inclinations' as more general dispositions or tastes, not necessarily tied to biological needs; and 'affects' as brief, forceful inner stirrings. Selecting carefully, we could construct a view, on Nietzsche's behalf, along roughly these lines. *Some* of the things described as 'drives' and 'instincts' by Nietzsche fit well into the category of the biological urge—sex, hunger, species-preservation (D 119; GS 1). In places, it seems they more or less constantly seek expression or satisfaction (D 119). Many 'inclinations' are general and nonbiological—such as those for luxury (GM III 8) or dialectics (D 48). *Some* affects look like candidates for brief, fiery emotional storms: anger, fear, hate, envy; sometimes, Nietzsche obviously uses 'affect' to indicate a seizure of excitement. In *The Birth of Tragedy*, for example, Apollo and Dionysus are 'drives' or, apparently equivalently, 'artistic drives of nature', operating through nature in general (BT 1, 2); 'fiery affects' are the cheap thrills Euripides gives his spectators (BT 12). (HH I 138 also uses 'affect' as violent emotional excitation.) Later, Nietzsche explicitly describes some affects as socialized, felt add-ons to instinctive activity, and his description of some 'tendencies' as 'counterinstinctive' (or of inclinations for 'unnatural' objects) might suggest a conceptual distinction between an instinct as biological or natural and a

tendency as socialized, occasionally nonnatural, disposition (see TI 'Untimely' 45; GM II 24; GS 80).

But this picture is drastically incomplete—for two reasons. The first is conceptual, relating to drives and instincts. The definition of 'drives and instincts' in the case of animals is unhelpful when applied to humans. Drives, recall, are the things that propel animals to behave quasi-rationally, quasi-purposively and quasi-expediently. But the salient feature of the animal, in contrast to the human, is that the non-'quasi-' versions are *lacking*: the human being building a home, unlike the bee, *does*, prima facie, have consciousness, reason, instruction, a plan, prior models, and so on. Given prima facie consciousness and so on, where do the drives come in? That, precisely, is what we want to know. Reminding ourselves of the biological context of these terms solves nothing. Indeed, it brings the problem into focus. Nietzsche probably knew this: his meditation notes that Strauss offers no remotely satisfactory *Darwinist* picture of how *humans* behave; later, he points out that moving from the nonconscious to the conscious case, from 'natural history' to human action, motivates the whole 'problem of consciousness', precisely because one could explain nonconscious behavior fully, but one could not apply the same explanations of nonconscious behavior to beings with consciousness (UM I 7; GS 354).

The second problem is textual. Put simply, the texts do not, as a whole, support the division just set out. We can see this from overlapping labeling and from specific examples. If the divisions were clear, then it would matter very much whether something was a drive (or instinct), affect, or tendency. Calling 'revenge' a drive suggests that it has some connection to our biological functioning and that it seeks expression. To say that revenge is an 'affect' suggests it is something we *feel*, briefly and stormily—not necessarily always seeking expression. To say that revenge is the *consequence* of an affect suggests, perhaps, that as a result of a seizure of excitement one undertakes to exact revenge. Yet, Nietzsche *does*, in different places, refer to revenge as a 'drive', 'affect', and consequence of an affect; elsewhere, 'revenge' is a nonunivocal term that can mask fear (or desire for honor) as the *real* motive. Yet, fear itself, in turn, is variously a drive, affect, or a 'feeling' that encourages a drive to act (revenge: TI 'Skirmishes' 6; GS 49; HH I 138; HH 'WS' 33; fear and similarly pity: GS 355; TI 'skirmishes' 45; D 109; TI 'Errors' 5). Anger can be a drive or an affect; it can also be a neutral state of affairs that a drive can interpret in order to express itself (D 119). These are not just isolated examples of overlapping labeling: *all* the typical affects I listed above are, at some point, counted as drives and/or instincts too, as are 'hope' and 'love'; furthermore, there is no sense that Nietzsche sees himself as changing his mind when moving from one to another.

As for specific examples, a closer look at what each term denotes also rules out neat interpretation according to the division sketched above. Nietzsche's drive-and-instinct pantheon is extraordinarily vast. As for drives, references to hunger are actually very rare. So often presented as a typical or intuitive drive for us to work with in accounts of Nietzsche's 'theory', I find it only once in his published works, when Nietzsche is specifically telling us that, qua drive, it is *atypical* (D 119; cf. Clark and Dudrick 2012: 167; Katsafanas 2013; Gemes 2013). Thus, in addition to conventional drives, we also find four other

types: (1) emotions or feelings (anger, fear, hope, hate, pity); (2) character traits (stubbornness, jest, combativeness, benevolence, tenderness); (3) social hostility (murder, thievery, revenge, appropriation, craving for power) and the opposite (submission, obedience, benevolence); (4) mental activity (reflection, recognition, doubting, nay-saying, holding-back; analyzing, researching); (5) miscellaneous ('drive to posit your ideal'; drives for certain specific kinds of food, for 'gentle sunlight and little talking'). The instincts mentioned by Nietzsche map well onto these categories, too: (1) pity, fear; (2) falseness with good conscience; delight in pretence; weariness; (3) enmity, gruesomeness, revenge, assaulting, and destroying; also obedience, subjection, and subservience; (4) intellectuality, understanding, and guessing right; (5) purity, the 'person-instinct'.

I have already suggested that the biological model of drive or instinct, based on the animal case, tells us very little about humans. But note that many 'drives' certainly do not obviously relate to physiological processes, do not seem standard features of every human, let alone nonhuman, animal, are not the sorts of things that we'd be inclined to say are naturally seeking expression. (Such lists, therefore, help caution against 'naturalistic' accounts of the drives if they suggest such properties. An adequate discussion would digress, but, taken as a whole, I can find nothing *distinctively* 'naturalistic' about the drive pantheon or Nietzsche's approach to constructing it. Richardson [2009], Gemes [2013], and Leiter [2009] favor the term in discussing drives.) While instinct *can* indicate something nonconscious, some of these examples of drives and instincts appear to *demand* or *presuppose* something like conscious reflection: mulling things over, intellectuality, researching, reflecting. And one can have a *Hang* for dialectics or for the *vita contemplativa*. When we take into account what Nietzsche tells us, affects obviously stretch well beyond brief, exhausting 'seizures of excitement' or even 'ways in which we feel' (Janaway 2009: 52) to incorporate desires and demands; indeed, some affects have 'wills' of their own (BGE 117).

Commentators who offer accounts of Nietzsche's drives rarely present the enormous variety of things he labels as drives, instincts, affects, and tendencies. Once this variety is laid out, it is harder to understand why we should expect to find anything like an account of what ties them all together. Certainly, this takes us a long way from the conventional distinctions for which, as discussed, we can find some support in Nietzsche's texts. Indeed, faced with such slippery labeling, we might be tempted to conclude that Nietzsche used these terms in a fluid manner, such that whenever he said the one he might as well have said the other. Clearly, there are aphorisms in which some two of these terms are used interchangeably (GS 1; BGE 36). Some scholars therefore implicitly present 'drive' and 'affect' or 'tendency' as interchangeable (May [2009: 90]; Katsafanas [2013: 747] on BGE 230 quotes from Nietzsche's description of a specific *Hang* to support his own general characterization of 'the drives'). But this allows Nietzsche too much, too quickly, because it suggests that he merely has two (or more) labels for one concept. 'Interchangeability' in one aphorism does not allow for interchangeability in all: otherwise, instinct and tendency would be interchangeable, and we would not find counterinstinctive tendencies; or affect and instinct would be interchangeable, and we would not find affects conceptually opposed to instincts, the former socially

grafted onto the latter (GM II 24; TI 'Untimely' 45). Furthermore, Nietzsche sometimes uses 'affects' interchangeably with 'feelings', though a drive, he suggests, is characterized by significantly more than what one feels (D 34; KSA 7: 364–5; see also D 115, GS 333; both are discussed below).

This discussion of Nietzsche's terminology and the variety of referents for 'drive' (etc.) shows the difficulty of finding either neat conceptual distinctions or a common point of reference. For all that, though, there may be some subset of uses of 'drive', 'instinct', 'affect', and 'tendency' that picks out some object that plays a substantive role in Nietzsche's understanding of mind. In the rest of this paper, I shall use 'drive' to name this. But I shall include passages that use all four of the labels under discussion, because each sometimes picks out something that performs the role of a drive in the coarse picture given above. There are costs to this way of proceeding: in particular, it suggests a unity where there is obviously disunity. But to answer our remaining two questions we shall need to have a term at hand.

2. Epistemology

Any attack on what we can know about actions and their causes would undermine the Socratic picture, (4) and (5). Nietzsche launches such attacks from D to TI (D 115, 116; GS 335; TI 'Errors' 3). The aphorism entitled 'the unknown world of the "subject"' claims that 'all actions are essentially unknown' (D 116). Moreover, GS 335 and GS 354 both assert that every human action is unique and unrepeatable, concluding that conceptual generalizations about action are inherently suspect; the former has every act as 'unknowable', the latter has conscious beliefs about actions as necessarily false. Nietzsche also warns us, specifically, how little one can know about *drives*, '*however* far a man may go in self-knowledge' (D 119, my emphasis). Furthermore, he gives us particular theoretical reasons regarding knowledge of drives. The battle of the drives that motivates our actions is said to be at a nonconscious (at least, not-fully-conscious) level (GS 111, D 129, GS 333). Of any one action, then, it is going to be difficult to say which drive or drives were primary. (Unless, as Nietzsche sometimes thinks, one drive is *always* the primary drive. GS 1 has species survival; GS 349 has 'expansion of power'; *D* has no fundamental drive. Nietzsche's occasional hunt for a primary or deepest drive is a further complicating factor for anyone seeking theory of drives from D onward. On power, see below.) It is difficult, even, to know which drives we have in the first place—hence discussions of the sort: is there *really* a drive to love or to knowledge? Nietzsche changes his mind about both (love: D 109; TI 'skirmishes' 40; knowledge: BGE 6; GS 110). But Nietzsche's epistemic skepticism goes further and turns into a skepticism about drive terms and their referents. In D 115, he contrasts the real, inner drives with what breaks through to consciousness: the breakthrough occurs when drives are in heightened states; in their milder or middling degrees we don't notice them. This has important implications for our *concepts* of, say, anger or pity, because we only name and conceptualize the excessive states—a bit like naming and conceptualizing an iceberg in terms of its tip. 'Anger', we saw, is variously

called a 'drive', an 'affect', and a state of affairs that may be interpreted by a drive. But, now, the drive we pick out by the label 'the anger drive' does not work very much like what we call 'anger': it is not just that you are angry more often than you think.

Now, the ineffectiveness of our names and our concepts for the drives, or the necessary falsification arising from conceptualizing every unique action, certainly stands in tension with the highly specific examples of drives (above). More important, though, the epistemic skepticism clashes with the philosophical *use* to which Nietzsche frequently puts his drives, which goes well beyond challenging the Socratic picture: (1) 'French' moralizing, that is, claims of the form: when you think you're acting out of pity (or selflessness, or benevolence), your motivation is *really* the gratification of some less noble drive (D 422, GS 49, GS 118); (2) comparative moral group psychology (D 38, D 78, GS 115–6, BGE 201, A 38): for example, there are two different names for the very same drive, according to whether the group in question is Greek or Christian; (3) drive biographies: explanations of the actions or characters of particular famous people, groups, or historical peoples by reference to their drives (D 245, 553, GS 99, 361, 368, BGE 200, 271, GM III 8, TI 'Skirmishes' 6, 45, *The Wagner Case*, 8 (KSA 6: 29–32); see Stern [2009] on Nietzschean freedom and drive biographies; epistemic concerns apply there, too).

I know of no attempt at an explanation anywhere in the literature—Nietzsche's or secondary—as to how Nietzsche can both hold that drive activity is in great part nonconscious, unknowable to individuals, and necessarily poorly conceptualized *and* claim intricate knowledge of the workings of the drives of others, to the extent that he can describe them in detail (as listed above) and make claims of these three kinds. Nor do I see how such an account is possible. To be sure, Nietzsche refers to himself as a 'psychologist without equal' (EH 'Books' 5). But using this quotation to endorse the hunt for a consistent, philosophically rigorous Nietzschean psychology is unwarranted (cf. Katsafanas 2013: 728): not only do EH's self-descriptions flirt notoriously with the megalomaniacal (Sloterdijk 2013: 48–50; Huenemann 2013: 77–8), more important, Nietzsche immediately gives an example of his great 'psychology'—a poetic encomium to Dionysus that has nothing to do with drives, consistency, or theory (EH 'Books' 6). Moreover, we now expect psychologists to tell us *how* they reach their conclusions: note TI's claim that 'psychologists' should not set out deliberately to observe people but, instead, should trust their 'instincts'—a demand that rules out a key methodological premise of empirical psychology at least (TI 'skirmishes' 7). In the next section, we look at ways in which drives relate to conscious deliberations. The more complex the theories provided to explain how drives work, the more pressing the epistemological problem becomes. Scholars who claim that Nietzsche's positive ideal is a matter of a particular, complex combination or configuration of the drives—for example, a 'maximally free and autonomous self' who has 'the maximum number and diversity of drives, each of them maximally powerful', hierarchically and aesthetically organized by a single idea (May 2009: 94)—owe us an account of how Nietzsche thinks we or anybody else could know that such a thing had been achieved (similarly for Gemes 2013 and Anderson 2013).

3. Drives and Conscious Deliberation

On the Socratic picture, conscious deliberations and decisions are fully causally efficacious. For Nietzsche, by contrast, *drives* are causally efficacious. There are many ways of spelling out this claim. Consider four general positions: 'overrule', 'manipulate', 'epiphenomenalist', and 'reductive'. Each admits of many versions, and some of these versions are incompatible with others. The four positions have two things in common: they challenge the Socratic picture, and at some point Nietzsche advocates each of them.

1. 'Overrule': The drives overrule our conscious deliberations and decisions.
2. 'Manipulate': The drives manipulate our conscious deliberations and decisions.
3. 'Epiphenomenalist': Only the drives are causally efficacious; our conscious deliberations and decisions are epiphenomenal.
4. 'Reductive': There aren't conscious deliberations and decisions, as we typically understand them, *at all*; there are only drives.

'Overrule' posits drives and deliberations as different things, which frequently oppose one another, just as, in the Socratic picture, there might be a tension between (anachronistically speaking) Socrates's survival instincts and his rational deliberations. But the balance of power has shifted. 'Overrule' admits of three distinct versions: (1) *strong*: deliberations always overruled; (2) *weak*: deliberations sometimes/often overruled, sometimes not; (3) *epistemic*: we never know if deliberations are overruled. ('*Weak overrule*' must be sufficiently strong that Socrates is wrong.) 'Epistemic overrule' fits well with Nietzsche's epistemic skepticism though that in turn, as discussed, does not fit with his other philosophical aims. 'Strong' is incompatible with both 'weak' and 'epistemic' (i.e., if conscious deliberations are *always* overruled, then it's not true that we never know).

The long analysis of D 129 combines 'weak' and 'epistemic'. (Other tentative versions of 'overrule' are less fleshed out; see BGE 158; GS 335.) Here, Nietzsche sets out the 'Socratic' picture: using 'reflective consciousness' to weigh up possible consequences, we 'often honestly torment ourselves' about the various things our deliberations might cause. Then we choose and believe that we understand our 'motive'. In fact, unknown and ill-known 'motives' (including some 'affect' that happens to 'leap in') battle it out—'something quite invisible to us of which we would be quite unconscious.' But, here, Nietzsche is *very clear* that the conscious deliberations are '*one* very essential motive in the battle line'. We might say: the 'Socratic' sees the deliberative battle as the 'final'; in fact, it is the 'semi-final', whose winner goes through to play several other, competing 'motives' in the 'final', of which consciousness learns only the result. This expressly allows for deliberations sometimes making the difference; often, though, they get overruled (outplayed) by other participants in the nonconscious battle.

'Manipulate' blurs the lines of the opposition suggested in 'overrule': it's not that your deliberations are overruled by your drives; rather, your deliberations are

'shot through' with drive activity, hence calling into question a 'drive-independent' position for the drives to overrule. Note the same three variations: (1) 'strong' (always); (2) 'weak' (sometimes, sometimes not); (3) 'epistemic' (we never know). Note, too, the same clashes: strong with weak, strong with epistemic.

Nietzsche's 'manipulate' claims do not typically allow us to choose between these three variations. But his writing is packed with them: the basic drive of GS 1 can be 'surrounded by a whole retinue of reasons [*Gründen*]' when it wishes and can even appear 'as Reason [*Vernunft*] itself'; Eleatics are criticized for 'deny[ing] the force of the drives in knowledge and generally [conceiving] reason as free, self-originated activity' (GS 110). The 'intellect', which seeks to control a drive, is *always* the 'blind tool of another drive' (D 109). As common are specific characterizations of a drive doing manipulative things: manipulating everyday experiences in accordance with its desires or 'seizing an event as its prey' (D 119); 'try[ing] to make us forget' about itself (GS 1); forcing its 'one-sided view' upon the situation (GS 333); ensuring that a person (consciously) believes something by keeping other things unconscious (TI 'untimely' 42). (See also BGE 6; BGE 59. Indeed, GS 333 has nonconscious drives as subjects of *verstehen*—they 'understand'.)

Consider, in particular, the discussion of dreaming at TI 'Errors' 4. The dream scenario is as follows (details added for clarity): a cannon shot is fired near the sleeping, dreaming Fritz; one of Fritz's drives intercepts the incoming sound and retrospectively furnishes it with a fake cause within Fritz's dream: now, in the dream, but *before* hearing the shot, Fritz is a brave medical orderly at the battle of Wörth; only once the cannon shot has been given an interpretation by the drive (i.e., it now makes sense that Fritz would hear a cannon, because he is on a battlefield) does the drive 'permit' it to enter Fritz's consciousness, because now it makes sense to Fritz as the *effect* of faked, battle causes within the dream.[3] For Fritz, Fritz hears the cannon because of his brave actions at Wörth; in fact, his brave actions at Wörth are caused by the cannon shot. Nietzsche then announces that we 'do just the same thing' when waking. This probably suggests 'strong manipulation': we have no direct access to perceptions that haven't been furnished, by drives, with causes. (Perhaps: drives have the option to manipulate all perceptions.) Certainly, this sits uncomfortably with the realm of drive-free (though 'overrule-able') deliberation suggested in D 129's 'overrule'. (HH I 13 and D 119 give weaker analyses of the same basic dream phenomenon. D 119 has waking responses 'essentially' the same as dreaming ones, if weaker. I need not choose which is authoritative, because my point is that this is typical of the difficulty of pinning Nietzsche to any position. But note that Leiter [2009] analyses TI in favor, ultimately, of a strong view, without treating the D or HH counterparts. Katsafanas [2013: 742] analyzes the weaker D 119 to support his weaker interpretation, but does not acknowledge that TI is stronger.)

---

3 Ken Gemes has rightly pointed out that Nietzsche does not explicitly say that the *drive* creates the interpretation itself. But Nietzsche does refer to a 'causes drive' ['*Ursachentrieb*'] which only permits the sensation to appear in consciousness once the sensation has been furnished with a cause. A natural interpretation has the causes drive furnishing the cause. Hollingdale agrees, choosing 'cause-creating drive' for clarity.

'Epiphenomenal' still has conscious deliberations and drives as different things. But now the former are rendered epiphenomenal. Indeed, D 124 suggests that all apparent instances of alignment between deliberation and action are mere coincidences, in which what we planned to do happens to line up with what we actually do: the person who says 'I will' and then acts accordingly is like the person who gets up before dawn and says 'I will that the sun shall rise' and thinks, when it does, that he has caused it to rise (see also D 130). Similarly, in TI, 'the will no longer moves anything, consequently explains anything—it merely accompanies events, it can also be absent.' This clearly rules out 'willing' as both a necessary and a sufficient condition for action. Nietzsche's conclusion: 'there are no mental causes whatsoever' (TI VI 3. In context, it's clear that 'will' is being used, here, to mean conscious thoughts and deliberations about which action to take.)

'Reductive' is primarily negative: it does not allow for drive activity and conscious deliberation as two separate realms. On the 'reductive' account, to contrast conscious deliberation with drive activity is to make a category error (rather as, in a classic example, the child might mistake the 'running' of a watch for something that exists and can be pointed to independently of the mechanisms of the watch): in fact, loosely put, deliberations are somehow made of drive. 'Reductive' positions are proposed, for example, in the suggestion that 'thinking is only the relationship of the drives to one another' (BGE 36, my translation), that the 'soul' is really a 'social structure of the drives and affects' (BGE 12, my translation), that our mistake is thinking that 'the understanding [*intelligere*] is . . . something essentially opposed to the instincts, when in fact *it is only a certain behaviour of the drives towards one another*' (GS 333). 'Reductive', too, are some of Nietzsche's most radical claims about the conscious 'ego [*Ich*]': it has become 'a fable, a fiction, a play on words' (TI 'Errors' 3).

An initial overview of these positions on drives and deliberations doesn't leave the prospects for a unified Nietzschean position in particularly good shape. Specifically, there are two central problems that, I want to suggest, we cannot resolve. Indeed, when we look into them more closely, the problems only deepen. The first problem is associated with Nietzsche's epiphenomenalist claims; the second is associated with his reductive claims. In the secondary literature, neither of these problems has been set out with sufficient clarity to allow for a direct solution. As to the first, two strategies present themselves. As to the second problem, even where it is recognized, the scale of the problem is hardly acknowledged. We look at each in turn.

4. The Problem of Epiphenomenalism

To begin with, the epiphenomenal position looks opposed at least to anything but the 'strong' versions of 'overrule' or 'manipulate'. It is flatly incompatible, of course, with D 129's 'overrule': compare the latter's 'one very essential motive' with TI's 'moving nothing' (TI VI 3 also has the notion of 'motive' *itself* as an 'error'). Though they are not inconsistent and though some may see them as mutually

reinforcing, it is decidedly odd to put 'strong manipulate' with 'epiphenomenal' accounts: frankly, why would the drives *bother* to manipulate something that caused nothing at all? (Leiter seems to put these two positions together [2009: 120–23].) If I tell you that the court just set the mafia boss free, but the trial was completely manipulated by the mafia, I think you would be right in assuming that the court's judgment itself is necessary or sufficient for the boss's freedom. Otherwise, why did the mafia bother with the manipulation? Recall, therefore, TI's explicit denial of conscious deliberations as necessary or sufficient conditions. Two strategies from the secondary literature present themselves at this point. Both are problematic.

4.1. First Strategy: Denial of Stronger or Weaker Claims

For many, Nietzsche just can't mean the 'stronger' things he says, in other words, the claims that look hardest to defend, philosophically: 'strong' overrule and manipulate and epiphenomenal (often not distinguished) (Leiter 2002: 92; Williams 1995: 70; Young 2010: 305; Gemes 2013: 563). Evidently, pursuing this requires setting aside some published statements by Nietzsche. More troubling, it requires ignoring what he himself *says* about the relationship between his stronger and weaker claims. In A 39, for example, Nietzsche *begins* with a 'weak overrule' claim: 'conscious states of affairs', such as believing, are weak motives when compared with the power of the instincts; then he moves to a stronger claim: '*more strictly speaking*, the whole concept of mental causality is false' (A 39, my emphasis; D 109 also begins 'weak' and then ends exclusively 'strong'). This suggests, if anything, the very opposite: that the language about deliberation being weak-but-partially-effective should ultimately be ignored; certainly, it is premature to dismiss the strong claims as hyperbole (cf. Gemes 2013: 563, fn. 19) if they are prefaced with 'more strictly speaking'.

A similar approach—denial of many strong claims—is taken (implicitly) by Katsafanas in his account of the drives, which is an extended 'weak manipulate' interpretation. Despite claiming to present an account of Nietzsche that 'does not require us to discount any published passages' (2013: 750), Katsafanas dismisses Nietzsche's epiphenomenalist claims on the grounds of some anti-epiphenomenal textual evidence and because they are philosophically implausible (2013: 749). (Both points are well taken; still, epiphenomenalism is sometimes Nietzsche's published view.) Within 'manipulate' itself, though, Katsafanas seeks to avoid the more philosophically implausible suggestions found in Nietzsche's descriptions: namely, that drives are homunculi—intentional agents that have independent perceptual capacities, desires, the ability to feel, make commands, obey, and so on (2013: 729–31). On Katsafanas's view, 'drives are dispositions that induce evaluative orientations' (2013: 740): by manipulating our conscious deliberations, drives mess with our beliefs, our affects, and our sense of what is appropriate for an apparently given (but in fact manipulated) situation. But (he claims) philosophically problematic, homunculi-type commitments are lacking: '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' (2013: 744).

Now it is doubtful, first of all, whether Katsafanas's own account actually manages to avoid the excesses of the homunculi claims. He frequently speaks of drives 'making . . . salient' what we would otherwise ignore: the hunger drive, for example, makes salient all the food sights and food smells that are anywhere near me (2013: 740–43). If all this means is that when I am hungry, I tend to notice food, then we are clearly a long way from Nietzsche's intentional drive language, which Katsafanas himself acknowledges and tries to account for. 'Making salient', though, can also be an intentional action, just as much, say, as 'giving orders' (which Katsafanas takes to be problematic): when I write a Nietzsche quotation in italics in a paper, I make it salient to you; obviously, this requires prior perceptual 'access' to the materials and a sense of how you will respond to them. (One could say: 'the italics made the quotation salient', hence avoiding an 'intentional' formulation. But drives are in no way analogous to italics.) Certainly, then, we are owed, on Katsafanas's terms, a better account of 'making salient' that doesn't turn drives into the homunculi he wants to avoid. Should Katsafanas be able to provide such an account, it would merely demonstrate how far he has come from *Nietzsche*, who (sometimes) evidently takes drives to have such independent capacities when they make things salient. First, consider his dream example (above) in which the relevant drive holds some sensation (in that example, the cannon shot) back from our consciousness until a cause for it has been fabricated. This clearly asserts two stages of perceptual awareness. Elsewhere, he gets as explicit as he possibly can that drives have independent consciousnesses from each other and from their host: 'I think these drives which fight each other here [i.e., 'unconscious and unfelt' by the person] know very well how to make themselves felt by and how to hurt *each other*' (GS 333, Nietzsche's emphasis).

Katsafanas also offers two exegetical arguments supporting his dismissal of homunculi claims: first, Nietzsche 'vociferously argues against the superfluous positing of subjects' (2013: 728) that homunculi readings would entail; second, Nietzsche wants to 'transform' or 'rethink' selfhood, whereas positing more minds entails more of the same, not radical rethinking (2013: 731). No textual support is offered for the first claim; indeed, Nietzsche sometimes *advocates* a 'subject-manifold' (BGE 12). Using TI 'Errors' 3 to support the second claim is misleading. In this passage, as in TI 'Reason' 5, Nietzsche claims that our concepts of causation and thinghood *tout court* come from our mistaken construction of ego as both thing and cause. It's not a call to rethink or transform the subject, nor is it an objection to *superfluous* positing of subjects. *Any* positing of subject *or* thing *or* cause is entangled in error (so Nietzsche claims). It would be hard to find an account of the drives that does not appeal to at least some of those notions and, unsurprisingly, neither Nietzsche nor Katsafanas offers one. As regards thought and action, the specific target is the Socratic picture. A many-drives-as-warring-persons account *would* challenge this, since for any *human* action there would no longer be a single, controlling mind. Specifically, the many-minds model challenges at least (1), (3), (4), and (5) of the Socratic picture. This leaves Nietzsche with a philosophically suspect position, but one that fits with his anti-Socratic aims. (This is not to suggest

that Nietzsche consistently adopts a many-minds model: he does not—see Gardner [forthcoming].)

An equal and opposite response to denying the stronger claims would be to deny the force of the weaker ones, bypassing suggestions that conscious deliberation can be causally effective in the actions we take. Leiter (2009: 118–23), quoting D 129 and arguing, ultimately, in favor of a version of epiphenomenalism, does not discuss the passage in which Nietzsche states that deliberations can be very powerful motives within the nonconscious battle. (Leiter's formulation of epiphenomenalism is qualified in various places [2009: 111, 119–20], but it is difficult to see any epiphenomenalism worthy of the name sitting with D 129.)

4.2. Second Strategy: Divide and Conquer

A more intriguing position divides the apparently contradictory claims into different philosophical categories. On this approach, although some conflict is acknowledged, the positions are made coherent by being placed in a broader Nietzschean framework: claims of the form 'A' are part of one project, but claims of the form 'not-A' are part of another. For Gardner (2009), for example, strong claims are part of Nietzsche's 'theoretical' project; the 'practical' project requires weaker claims. Gardner's Nietzsche is therefore sometimes doing one comprehensible though not, he convincingly argues, philosophically coherent, *theoretical* thing and sometimes doing a completely different *practical* thing. (Janaway [2009: 59] seems to echo this division: he speaks of the conflict between Nietzsche's 'official' [strong] position and his practical desire for a unified self.)

Even this, though, underplays the deep contradictions within Nietzsche's remarks: there is plenty on the '*theoretical*' side that contradicts what Gardner takes as Nietzsche's theoretical view. Some of these contradictions we have already seen in various claims about causally effective conscious deliberations that seem unconnected with any particular 'practical' project (D 129; GS 335; D 115; BGE 3 claims not that *all* of conscious thought can be ascribed to the activity of the instincts, but only 'the greatest part'). But we should also note that Nietzsche gives his own 'theoretical' *arguments* against the strong claims. For instance, GS 354 begins by making the familiar point that much of what we take conscious activity to be causing is done perfectly well by nonconscious animals. But, Nietzsche argues, we just *cannot* move from that to positing a consciousness that serves no purpose, because it would be absurd for animals to develop with such a complicated but useless mechanism. In explicating the model (from contemporary empirical psychology) that Leiter takes Nietzsche to have anticipated, Leiter (2009: 122) compares deliberation and action to thinking that one is playing a computer game when in fact one is watching a film of standard footage, in which the film just happens to correspond to the buttons one presses. This helpfully sets out a position Nietzsche sometimes takes. But what, Nietzsche is asking in GS 354, would be the point of an animal evolving in such a peculiar way? His conclusion is that one simply can't posit such a thing. This is the sort of consideration that sometimes leads Nietzsche to allow for a small chink of causally effective conscious thought.

Thus, if Gardner is correct, not only does Nietzsche not mean what he sometimes claims about causal effectiveness; he has solved or forgotten about his own perfectly valid objections to his very position. (Strictly speaking, GS 354's focus is not consciousness-to-action but rather action-to-consciousness. This is problematic for Gardner's 'theoretical' view and also indirectly problematic for epiphenomenalist views of deliberation/action that, along the line Nietzsche suggests, would need to explain *why* we go to the trouble of developing epiphenomenal deliberations.)

But by the same token, we cannot see GS 354 as Nietzsche's final word on this question. Richardson (2009), who attempts this, takes GS 354 as the platform for a different way of combining Nietzsche's apparently contradictory claims according to which drives are individual, presocialized animal instincts, whereas consciousness and language are a later, socialized, and occasionally causally effective response to the animal drives. Richardson also, in the end, denies some stronger claims, since his account is not epiphenomenal: consciousness holds back or prevents drives from bringing about actions. But note, first, that TI's and A's denials of mental causation, quoted above, come after GS 354 chronologically and insistently overrule it. Second, the sorts of things that Nietzsche tells us count as drives are frequently neither presocial nor prelinguistic: consider the drive for 'little conversation' or 'benevolence'. Finally, the idea that consciousness holds back or prevents drives from being effective ignores the very drives Nietzsche mentions that evidently serve just the purpose Richardson ascribes to consciousness in *opposition* to the drives: holding back, not reacting immediately, weighing things up, mulling things over. There is, then, no obvious way to solve the problem of epiphenomenalism.

5. The Reduction Problem

'Reductive' claims—however they are understood—pose their own problems that are rarely if ever acknowledged in the literature. As we shall see, there is no unproblematic way of understanding how, for Nietzsche, conscious deliberation reduces to drive activity. For one thing, he sometimes seems to warn against the coherence of even asking whether consciousness is 'real' (BGE 34). For another, many drives, as already indicated, are *identified* in terms that suggest the sort of mental activity that is supposed to be 'reduced' (see above).

That aside, to say that 'A reduces to B' leaves it open what we want to say about the causal properties of A. Here are three versions. Benign reduction: we reductively redescribe your 'punch' in scientific terms; still, your punch caused my bruise. Instrumental reduction: A reduces to B such that A does not strictly exist and causes nothing; nonetheless, speaking about A as causal is a very useful 'as-if' activity (e.g., centers of gravity). Eliminativism: B accounts for everything that A attempts (poorly) to account for and more besides; we ought not to talk in terms of A *at all*.

It is implausible that Nietzsche thought he was advocating benign reduction: his reduction claims are corrective, criticizing the Socratic picture. Eliminativism

presents other concerns. It denies a premise of overrule, manipulate, and epiphenomenal: that we must think of the drives as something conceptually *opposed* to conscious deliberation. This is obvious enough for 'overrule' and for the weak 'manipulate'. The strong 'manipulate' suggests that deliberating is something *conceptually distinct* from drive activity: the two could—at least in theory—come apart. And an epiphenomenal, conscious 'deliberation' is still a conscious deliberation which, as causally inefficacious, contrasts with the causally efficacious activity of the drives. If Nietzsche's project really were to eliminate mind talk in favor of a superior drive talk, of course, then his deep skepticism about the possibility of knowledge of drives (see above) would be all the more perplexing. Finally, at least regarding the soul, Nietzsche is sometimes explicit that, in reducing it, he need not eliminate it altogether [*los zu werden*] (BGE 12).

There is some support for instrumental reduction, especially in the notebooks (see WLN 20--21; see also perhaps BGE 34). And it might help interpret a claim such as that conscious deliberations are overruled by drives: strictly speaking, the former reduce to the latter, but it's helpful to speak as if they are distinct. Yet this vastly underestimates how often Nietzsche's discussions manifestly treat drives and deliberations as conceptually *opposed* (e.g., BT 13, D 58, 221, 328; GS 11, 294; BGE 158, 284; GM I 10, III 9; TI 'Untimely' 7, 38). Certain religious codes, for example, are supposed to 'push back' [*zurückzudrängen*] consciousness, so that the instincts are left to a state of 'perfect automatism', that is, so that consciousness cannot interfere with the instincts (A 57). But two further passages deserve more attention. First, BGE 191 *criticizes* Socrates for trying to 'talk Reason into' [*die Vernunft überreden*] working for the instincts; yet, what Socrates is doing (in BGE 191) looks identical to *Nietzsche's* claims in GS 333 or BGE 3 about how things work anyway. Second, D 560 asks us to 'dispose of our drives like a gardener', in other words to think about how much 'anger, pity, curiosity, vanity' (all listed as drives) we would like and in what relation we should like to stand to them and they to each other. Only *one* of the *many* options is for 'us' to let 'them' (i.e., the drives) fight it out among themselves, which looks for all the world like what Nietzsche elsewhere says that they do anyway, whatever we 'choose' to do about it. There is not a hint, in D 560, either of the unknowability of what drives are nor of the drives having a deep-seated control over what we do and how we behave toward them. To hold this together with some other claims by Nietzsche, the gardener carefully pruning the roses in the English fashion would have only ever seen the tip of a rose petal, would be the 'blind tool' of the lilies and would, himself, be made of carnations.

Passages like these reveal the challenge of pursuing instrumental reduction in Nietzsche. First, *drives* themselves are characterized by their *opposition* to consciousness. Much *drive* talk must therefore also be 'instrumental' (or plain inconsistent), proceeding as if drives are opposed to consciousness, when in fact it is composed of drives. Second, it merely pushes back to the 'instrumental' level the same set of interpretative problems above: if it matters whether ('as if') 'conscious deliberations' are ('as if') 'causally effective'—if answering that question has philosophical significance—then Nietzsche fails to give an answer.

Note, finally, that practically no position I set out here is confined to one Nietzschean 'era' (from D onward): skeptical, epiphenomenal, antireductive, and manipulative claims are found from D to TI; reductive, from GS (I–IV) to TI; anti-epiphenomenal, from D to GS (V). (Overrule, exceptionally, is most explicit in D; its anti-epiphenomenal, antireductive import continues, as we have seen.) Hence, it is unlikely that restricting one's focus to one Nietzschean book, year, or era can solve these problems, and there is no such attempt in the literature. While it is undeniable that Nietzsche's ideas develop throughout the period under discussion, it is unlikely that these developments—themselves often ambiguous and controversial—clarify rather than further complicate Nietzsche's discussions. A case in point is the will to power, emergent after GS (I–IV), which plays at least two apparently inconsistent Nietzschean roles: (1) explaining *all* biological activity (BGE 13, GS 349); (2) selectively indicating the praiseworthy, hence *not* a feature of all biological activity (A 6). I agree with Young (2010: ch. 26) that the latter is dominant and less problematic. I cannot see either resolving the challenges I have presented, since both are prima facie compatible with each position. To say that my 'truly basic life-instinct' is will to power does not distinguish between conscious deliberations as overruled by, manipulated by, epiphenomenal to, or reducing to such an instinct. Similarly, stating that I am praiseworthy when motivated by, for example, will to power leaves open epistemic concerns and the question of how such motivation interacts with conscious deliberation. (The best attempt at combining a theory of drives with will to power, Richardson [2013], rests upon Richardson [2009], which I criticize above. On will to power, including an independent discussion of Richardson [2013], see Stern [forthcoming].)

6. Conclusion

In the background lies the question of what we want from a philosopher like Nietzsche. His 'drive' terminology is confusing; his skepticism about drives and actions is as far-reaching as the knowledge he claims (and needs) to have about them; the relation between drives and conscious deliberations is expounded in an overwhelming variety of different and often contradictory ways. This much is clear: if 'success' in philosophical writing means the production of an exegetically plausible, unique, and philosophically robust theory—the sort of thing that many scholars have aimed to produce from him—then it must be acknowledged that Nietzsche failed in this case. On that assumption, one should abandon him as philosophical psychologist and/or produce one's own, Nietzsche-inspired theory (provided the clash with Nietzsche is acknowledged). Is there a different kind of 'success'? Nietzsche was interested in something that still interests us: challenging the Socratic model. It is philosophically fruitful to think with him as an experimenter and provoker who does not set himself the task of producing, expounding, and defending a particular theory, but who throws everything he can at his target. His efforts, laid out now in the broad daylight of their contradictions, offer us an insight into different modes of opposing Socratism *and* a profound

sense of discomfort with each. If, taken as a whole, this is a challenging, valuable, and illuminating way of thinking, then it can attain this value precisely when one *abandons* the hunt for the unique, final theory. The worry with the resultant theories is not merely that they clash with what Nietzsche says; they may also clash with what he was trying to *do*. I don't take myself to have established this last point. But I do think it worthy of consideration. Perhaps, to appreciate and criticize what Nietzsche is up to, we must stop reading him as though he is up to something else.

TOM STERN
UNIVERSITY COLLEGE LONDON
t.stern@ucl.ac.uk

References

NIETZSCHE'S WORKS

Nietzsche, F. *Sämtliche Werke, Kritische Studienausgabe in 15 Bänden* (KSA). Edited by Giorgio Colli and Mazzino Montinari. Berlin: De Gruyter, 1999.

Nietzsche, F. *The Anti-Christ* (A). In R. J. Hollingdale (trans.), *Twilight of the Idols and The Anti-Christ* (London: Penguin, 1968).

Nietzsche, F. *Beyond Good and Evil* (BGE). In W. Kaufmann (trans.), *Basic Writings of Nietzsche* (New York: The Modern Library, 2000).

Nietzsche, F. *The Birth of Tragedy* (BT). Translated by Ronald Speirs. Cambridge, UK: Cambridge University Press, 1999.

Nietzsche, F. *Daybreak* (D). Translated by R. J. Hollingdale. Cambridge, UK: Cambridge University Press, 1997.

Nietzsche, F. *Ecce Homo*. Translated by W. Kaufmann in *Basic Writings of Nietzsche*. New York: The Modern Library, 2000.

Nietzsche, F. *The Gay Science* (GS). Translated by Josef Nauckhoff. Cambridge, UK: Cambridge University Press, 2001.

Nietzsche, F. *Human, All Too Human* (HH). Translated by R. J. Hollingdale. Cambridge, UK: Cambridge University Press, 1996.

Nietzsche, F. *On the Genealogy of Morality* (GM). Translated by Carol Diethe. Cambridge, UK: Cambridge University Press, 1997.

Nietzsche, F. *Twilight of the Idols* (TI). In R. J. Hollingdale (trans.), *Twilight of the Idols and The Anti-Christ* (London: Penguin, 1968).

Nietzsche, F. *Untimely Meditations* (UM). Translated by R. J. Hollingdale. Cambridge, UK: Cambridge University Press, 1997.

Nietzsche, F. *Writings from the Late Notebooks* (WLN). Edited by R. Bittner. Cambridge, UK: Cambridge University Press, 2003.

WORKS BY OTHER AUTHORS

Anderson, R. L. (2013) 'Nietzsche on Autonomy'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 432–60.

Clark, M., and D. Dudrick. (2009) 'Nietzsche on the Will: An Analysis of BGE 19'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 247–68.

Clark, M., and D. Dudrick. (2012) *The Soul of Nietzsche's* Beyond Good and Evil. Cambridge, UK: Cambridge University Press.

Gardner, S. (2009) 'Nietzsche, the Self and the Disunity of Philosophical Reason'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 1–31.

Gardner, S. (Forthcoming) 'Nietzsche and Freud: The "I" and Its Drives'. In Maria Branco and João constâncio (eds.), *Nietzsche and the Contemporary Debate on the Self* (The Hague: Walter de Gruyter, forthcoming 2015).

Gemes, K. (2013) 'Life's Perspectives'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 553–75.

Harte, V. (2005) 'Conflicting Values in Plato's *Crito*'. In R. Kamtekar (ed.), *Euthyphro, Apology, and Crito: Critical Essays* (Lanham, MD: Rowman and Littlefield), 229–59.

Huenemann, C. (2013) 'Nietzsche's Illness'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 63–80.

James, W. ([1894] 1920) 'The Physical Basis of Emotion'. In R. B. Perry (ed.), *Collected Essays and Reviews*, (London: Longman, Green), 346–70.

Janaway, C. (2009) 'Autonomy, Affect and the Self in Nietzsche's Project of Genealogy'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 51–68.

Kant, I. (1996) *The Metaphysics of Morals*. In M. Gregor (trans.), *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Cambridge, UK: Cambridge University Press.

Kant, I. (1997) *Die Metaphysik der Sitten*. Stuttgart: Reclam.

Katsafanas, P. (2013) 'Nietzsche's Philosophical Psychology'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 726–55.

Leiter, B. (2002) *Nietzsche on Morality*. London: Routledge.

Leiter, B. (2009) 'Nietzsche's Theory of the Will'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 107–26.

May, S. (2009) 'Nihilism and the Free Self'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 89–106.

Plato. (1997) *Crito*. Translated by G. M. A. Grube. In J. M. Cooper (ed.), *Plato: Complete Works* (Indianapolis: Hackett), 37–48.

Plato. (1952) *Phaedo*. Translated by R. Hackforth. New York: Liberal Arts Press.

Richardson, J. (2009) 'Nietzsche's Freedoms'. In K. Gemes and S. May (eds.), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press), 127–49.

Richardson, J. (2013) 'Life's Ends'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 756–84.

Sloterdijk, P. (2013) *Nietzsche Apostle*. London: MIT Press.

Stern, T. (2009) 'Nietzsche, Freedom and Writing Lives'. *Arion*, 17 (1), 85–110.

Stern, T. (Forthcoming) 'Nietzsche's Ethics of Affirmation and Negation', working paper.

Strauss, D. (1895) *Der alte und der neue Glaube: Ein Bekenntnis*. Bonn: E Strauss.

Williams, B. (1995) *Making Sense of Humanity*. Cambridge, UK: Cambridge University Press.

Young, J. (2010) *Friedrich Nietzsche: A Philosophical Biography*. Cambridge, UK: Cambridge University Press.