

## Drives as Inverted Forms: Nietzsche's Correction of Socrates's Philosophical Psychology

*As pulsões como formas invertidas: a correção de Nietzsche à psicologia filosófica de Sócrates*

Brian LIGHTBODY  
Professor  
Chair Philosophy Department  
Brock University.  
E-mail: [blightbody@brocku.ca](mailto:blightbody@brocku.ca)

### ABSTRACT:

A recent paper by Tom Stern suggests that Socrates's philosophical psychology, which emphasizes rational reflection, is superior to Nietzsche's drive model when explaining human behavior. I argue that Stern's analysis is wrong on three fronts. First, the models share common, though inverted, features. Second, Stern fails to consider the role of Socrates's *daimon* when evaluating Socrates's philosophy of mind; third, Nietzsche's model is more warranted. In sum, Nietzsche's philosophical psychology is a correction of the Socratic.

### KEYWORDS:

Drives, Socrates, Nietzsche, *Daimon*, Elenchus, Abduction, Forms.

### RESUMO:

Um artigo recente de Tom Stern sugere que a psicologia filosófica de Sócrates, que enfatiza a reflexão racional, é superior ao modelo pulsional de Nietzsche ao explicar o comportamento humano. Argumento que a análise de Stern está errada em três frentes. Primeiro, os modelos compartilham características comuns, embora invertidas. Em segundo lugar, Stern não considera o papel do *daimon* de Sócrates ao avaliar a filosofia da mente de Sócrates; terceiro, o modelo de Nietzsche é mais justificado. Em suma, a psicologia filosófica de Nietzsche é uma correção do socrático.

### PALAVRAS-CHAVE:

Pulsões, Sócrates, Nietzsche, *Daimon*, Elenchus, Abdução, Formas.

### Introduction

Nietzsche's reflections on drives in such works as *Human all Too Human*, I. 32, *Daybreak* 109, 119, *Beyond Good and Evil* 9, 19, 36 and *On the Genealogy of Morals III*: 8), have given way to a perspicuous reconstruction of drive theory in the secondary literature. While it would be beyond the scope of this paper to define all of the positions on Nietzsche's view of drives in the scholarship, one can say that

drives motivate behaviors and thus are the bedrock for Nietzsche's system of values and psychology.<sup>1</sup> As Paul Katsafanas summarizes its role, the drive is, "Nietzsche's principal explanatory token within psychology."<sup>2</sup> According to the secondary literature, the statement accurately describes the role of drives in explaining agential choice and action.

Yet, as Tom Stern argues in his incisive "Against Nietzsche's 'Theory' of Drives," drives are explanatorily idle: as transformed animal instincts, drives are too crude to explain deliberation and therefore provide us with little motivation to reject the standard account of agential choice, which puts pride of place on the operations of reflective consideration when making a decision (e.g. evidence gathering, analysis, the weight of reasons to undertake some action as opposed to some other etc.). For clarity, I shall follow Stern and call the view of mind, which stipulates that reasons are the true engines for decision-making, the "Socratic model." The position is Socratic because Nietzsche, at least on the surface, offers drive theory as an alternative to this ancient template of mind, as evidenced by the following two passages in Nietzsche's *oeuvre*. The first is the 'Problem of Socrates' in *Twilight of the Idols*, especially section 11: "Is it necessary to go on to point out the error which lay in his faith in 'rationality at any cost?...The harshest daylight, rationality at any cost, life bright, cold, circumspect, conscious, without instinct, in opposition to the instincts, has itself been no more than a form of sickness.." (II: The Problem of Socrates, 11). The second is found in *Ecce Homo*, where Nietzsche reaffirms his initial diagnosis in *Twilight* stating that Socratism may be defined as "rationality at any price." It is, he continues, "a dangerous force that undermines life." (E.H. BT 1). From such quotations, it is tempting to think that the two psychological models are worlds apart and that one must choose between them. Indeed, Stern draws this very conclusion, arguing that Nietzschean "drive theory" is an inferior model of our collective psychology compared to the Socratic and should, therefore, be rejected.

Yet there are other places in Nietzsche's work where he acknowledges Socrates's influence. In the "Struggle Between Science and Wisdom" (1875) Nietzsche writes: "Simply to acknowledge the fact: Socrates is so close to me that I am almost continually fighting with him." (SSW 188). But in what manner is Socrates "close" to Nietzsche? Some, like Kaufmann, have argued that the similarities between the two are stylistic: *Ecce Homo*, Kaufmann avers, is Nietzsche's attempt "...to trump the matchless irony of the

---

<sup>1</sup> I write "minimally" because there are more controversial, robust views of drives in the literature. One such view is the "homunculi" drive position which stipulates that drives are like proto-agents. See Maudemarie Clark and David Dudrick's *The Soul of Nietzsche's Beyond Good and Evil*, (Cambridge, Cambridge University Press), 2012

<sup>2</sup> Paul Katsafanas, *The Nietzschean Self: Moral Psychology, Agency and the Unconscious*, New York: Oxford University Press, 2016, 74.

*Apology*.<sup>3</sup> No doubt Kaufmann has been proven right in that regard, but the closeness between the two philosophers goes well beyond mere style, or so I shall argue. There are, in fact, substantive similarities between the two philosophers, especially in regard to their notions of mind. Thus, I argue Stern is incorrect on two fronts. First, Socrates's model of human psychology is not anathema to Nietzsche's: the two have much in common. Specifically, both models employ "givens": abductive hypotheses. To wit, there seem to be underlying psychological structures that, although they cannot be directly investigated if assumed to be present, make sense of human behavior. Nevertheless, although both rely on assumptions, Nietzsche's psychological model is, in fact, the more parsimonious of the two and, therefore, is more justifiable, all things considered, in contrast to Stern's diagnosis.

Secondly and even more profoundly, both models possess identical, though inverted, components. Consider the following three aspects each model shares: 1) they reflect a ground that cannot be investigated directly. For Nietzsche, these are animal instincts, while for Socrates, these are Forms.

2) Drives *qua* mental are conscious urges underpinned by not fully cognizable biologically rooted motivations. They propel agents toward some object or activity that would satisfy the active drive. For Socrates, his *Daimon* restrains him from action or inaction. To summarize, drives, say go! They push a subject forward. In contrast, the *daimon* says stop! It pulls the subject back.

3) Both models have fixed channels for expressing their fundamental tokens of psychological currency; drives have definite aims connected to their cravings. The drive points out a feature of the environment to its physical carrier and urges the 'agent' to act or acquire the object to quench the drive's thirst. If the individual successfully satisfies her urge, another drive is queued up, and the process starts all over again. If not, the drive continues to exhibit its peculiar cravings and demands satisfaction. Similarly, for Socrates, our reasoning always reaches out to full expression via specific virtues.<sup>4</sup> In some cases, it is clear what virtue needs to be applied and how, while in others, the situation is more opaque and requires more careful consideration.<sup>5</sup>

---

<sup>3</sup> Walter Kaufmann, "Nietzsche's Admiration of Socrates" *Journal of the History of Ideas*, 1948 472-491.

<sup>4</sup> See Daniel T. Devereaux, "The Unity of the Virtues in Plato's Protagoras and Laches", *The Philosophical Review*, Vol. 101, No. 4 (Oct., 1992), pp. 765-789.

<sup>5</sup> The early Socratic dialogues are filled with self-righteous individuals who believe they have a clear sense of what they are doing and why they are acting virtuously. The most obvious example is *Euthyphro* who prosecutes his own father for murder and declares that what he is doing is pious. In his own words, He claims to know precisely about all such matters" (5a) But like all Socratic interlocutors, Euthyphro becomes befuddled after engaging in a dialectical exchange with the gadfly of Athens. On the other side, consider Socrates defense in the *Apology*. There Socrates is defending himself from three charges. He understands that if he continues to speak in his usual manner, namely, his method of elenchus, he could

I conclude that Nietzsche's account of drives is preferable to the Socratic model of mind because it is more illuminating and possibly more fruitful. As a naturalistic theory about the human condition, drives can potentially be verified (or falsified) by our best psychological sciences. Far from being antithetical, it is more correct to say that Nietzsche's drive theory marks a significant improvement over Socratic philosophical psychology. Indeed, drive theory is a correction of sorts to the theory of the Forms, and it is this correction that explains Nietzsche's "close to me" remark about Socrates noted above.

I : What are Nietzschean Drives?

Mattia Riccardi provides a five-fold sketch of drive theory in his insightful and recent book, *Nietzsche's Philosophical Psychology*. He argues that drives have five aspects: 1) They are biologically rooted; 2) They have definite aims; 3) There is a close causal relationship between drives and affects; 4) Drives evaluate, which is to say they make things in the world salient which are then deemed worthy of pursuit and 5) Drives are propulsive; they urge their physical carriers to pursue those things evaluated as good. It would be beyond the scope of this paper to articulate all 5 aspects of drive theory (Riccardi, 2021, Chapter 1 section 2). Thus, I will focus my attention on aspects 1, 2, and 5: the biological rootedness of drives, their fixed aims, and their propulsive properties. It is these three aspects that outline the animality of drives.

The first aspect claims that drives are grounded—in some capacity—in one's biological constitution.<sup>6</sup> John Richardson is the scholar who has done the most and, in my opinion, best work in stressing the physiological substrate that subtends drives. As Richardson puts it in his *Nietzsche's New Darwinism*, "I claim that we can't understand his views on our values without seeing first and precisely how he thinks we are animals with drives" (Richardson, 2004 11-12). This way of putting the matter will be a refrain to which I return. Moreover, Richardson makes a strong case for suggesting that one cannot provide a coherent view of drive theory without seriously considering what Nietzsche has to say about

---

very well be put to death. But at no point is he contrite; his moral and philosophical clarity is beyond reproach. He continues to philosophize despite whatever dangers he may face for doing so.

<sup>6</sup> Whether all drives are remnants of animal instincts and in what sense instincts are different than drives, is an open question. For example, in his recent book, *Nietzsche's Moral Psychology*, Mark Alfano makes a persuasive case for arguing that instincts are a subset of drives: "I now want to argue that instincts are a species of drive: instincts are innate drives." Mark Alfano, *Nietzsche's Moral Psychology*, (Cambridge University Press, 2019), 69. However, and in keeping with Nietzsche's naturalistic hypothesis, I will treat drives (*Trieb*s) and instincts (*Instinkts*) as co-referential: both are expressions of a physical need or a transformation thereof.

breeding (züchten)—a polysemous term for Nietzsche that includes inculturation and biological inheritance.

Richardson's task is truly Herculean; he attempts to provide a feasible account for explaining how humans became rational, moralistic animals. Under Richardson's narrative, drives remain the elemental conduit of lineage for Nietzsche, yet they can be exapted for different purposes. "Thus drives can retain an end, but this end need not be conscious; nor is the end non-revisable. In better moments he treats drives as designed by selection. They are so designed simply qua drives, in that organisms are crucially rendered fit by being equipped with plastic dispositions (drives) physical set ups with causal tendencies that are plastic toward certain ends" (Richardson, 2004, 40). In worse moments, which Richardson avows are far more often, Nietzsche treats drives as little homunculi or (little agents) a problem that we cannot address here, unfortunately.

It is crucial to think through the implications of this commitment to, shall we say, the bestial origin of drives. Indeed, it would be difficult not to conceive of drives as having an animal origin, at least when we view Nietzsche's middle to late works. Consider Nietzsche's famous passage of *Beyond Good and Evil*: "We must translate man back into nature" (Nietzsche, 2000, BGE 230) or the start of *On the Genealogy of Morals* where Nietzsche intimates that his entire genealogical project may be construed as providing a naturalistic account of how the promise-making animal man came to be (Nietzsche 2000, GM I: 1). The obvious physiological platforming of drives gives way to their mental representation. I now turn to the phenomenology of drives.

#### Aspect Two: Propulsion

Drives urge their bearers to initiate a series of steps to complete some action over some other. When we experience a drive, we feel compelled to express it. The idea here is that a subject performs some action because the subject is compelled to either *simpliciter*, deliberation is causally neutral, or a drive manipulates the subject to do so. Regardless of the determinism (e.g. soft or hard) one adopts in relation to drives, what is clear is that drives motivate action. From there, some (Leiter) suggest that the drive alone is responsible for action which is to say that deliberation is causally inefficacious, while others (notably Katsafanas) take a more nuanced approach, arguing that drives influence choice but not the other way around.<sup>7</sup> Presenting this softer position in greater detail, we can say that once a drive is

---

<sup>7</sup>"Each person, Leiter declares, has a fixed psycho-physical constitution, which defines him as a particular type of person."

Brian Leiter, "The Paradox of Fatalism and Self-Creation in Nietzsche," in *Willing and Nothingness: Schopenhauer as Nietzsche's Educator*, ed. Christopher Janaway (Oxford: Clarendon, 1998), 219 "This is the problem that Nietzsche's drive psychology raises: choice may control action, but agents do not control choice." Paul Katsafanas, "Nietzsche's Philosophical

activated, like the sex drive, we have the freedom to express and satisfy the drive through our actions. However, we do not have the power to stand back and negate the drive, as it were, through an act of deliberative intention. The drive is either satisfied by some external action or becomes internalized due to the inability of the agent to express it. Drives, minimally, then, seem to have some causal influence over a subject's action and initiate a call to action independently of reflection.

Aspect Three: Drives have fixed aims

Drives seek to express themselves but do so in designated channels. As Nietzsche puts it in *The Gay Science*: Drives are “a quantum of dammed up energy waiting to be used up somehow” (Nietzsche 1974, 360) Aspect two clarifies the expressive facet of this statement. In this paragraph, I explain the “somehow” component. According to scholars in the literature, drive expression has both an aim and an object. The aim of the drive is fixed, and for most scholars, the object of the aim may change. For example, the sex drive leads a subject to find a suitable mating partner. However, according to the majority of scholars, the object or in relation to the sex drive, the specific partner, is unessential. As Mark Alfano puts it: “a drive is a disposition to activate behavior of a particular type though not necessarily with respect to any particular intentional object” (Alfano 2019, 51). Alfano's example of the expression of the sex drive is constructive in enlarging this detail. He continues, “someone's sex drive might lead them to engage in sexual congress with one person today to perform a different sex act with a different person tomorrow or to masturbate alone the next day” (Alfano 2019, 51).

In an earlier article, Paul Katsafanas echoed this sentiment. Drives he argued do not aim at the achievement of some determinate state of affairs. To concretize this a bit more, Katsafanas claims that “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” (Katsafanas 2013, 752). Katsafanas' explication is necessarily convoluted because it is difficult to maintain that a drive's expression is fixed, yet its aim might change. To take animalistic drives such as the thirst, shelter, sex, and hunger drive; the drives are distinct (we understand how each drive may pursue objects for its expression), and yet the aim of these drives will differ following the things in the environment that might satisfy the drive.

The above two components are intimately connected. Earlier, I mentioned that the phenomenological experience of an urge to perform some activity or find an object when a drive takes

---

Psychology'. In K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press. 2013), 752

hold represents the drive's mental component, while the third aspect illuminates the external means available for the drive to pursue this craving. These two components are obviously connected in that the drive's capacity to make something salient for its bearer is so that the object can satisfy the yearning of the agent and, therefore, quench the drive's thirst, if only momentarily. The appetitive nature of the drive (i.e., its propulsive property) bridges the drive's biological platforming (i.e., animal instincts) and the environmental conditions that allow it to be expressed. In finding a suitable object or activity, the urge dissipates, allowing the next drive to be queued up, where the process starts all over again.

Keeping the picture of the animal base of drives in our sights is vital when thinking about their aim. The purpose of the division between the drives' aim and its eventuated expression is to view human behavior *sans* intentionality; drives are meant to pull back the curtain on what some philosophers call action theory. The object that satisfies the aim is inconsequential; an agent does not really choose it for its own sake. It is instead a means for the unconscious need to express the active drive. The purported purpose of Nietzsche's moral psychology is to reveal the true motives that underpin the reasons that supposedly subtend our intentional decisions.

The idea that drives and drives alone explain all that there is behind goal-directed behavior is well-supported in Nietzsche's *oeuvre*. I point to a representative passage from *Gay Science* 354 "The Problem of Consciousness": "The whole of life would be possible without its seeing itself as it were in a mirror: as in fact even at present the far greater part of our life still goes on without this mirroring, and even our thinking, feeling, volitional life as well, however painful this statement may sound to an older philosopher" (Nietzsche 1974, 354).

This statement has raised several perplexing questions in the secondary literature in relation to the nature of consciousness. Consider that if Nietzsche is not being facetious when he states that the whole of life would be possible without seeing itself as it were in the mirror, that is our lives would carry on pretty much as they are now without the advent of consciousness, then what is the purpose of consciousness generally, when it is, in the main, superfluous? Stern's question in relation to this passage can be summarized as "Can Nietzsche's model demonstrate that consciousness is unnecessary vis a vis understanding our actions?"

However, the argument runs the other way too. If consciousness is but a mirror for unconscious drives, how do we know the mirror accurately represents these fundamental psychological components? As noted by Katsafanas above, drives "...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 aspect of drives applies to material objects in the environment and abstract immaterial

entities alike. As Nietzsche admits in *Beyond Good and Evil*: “But anyone who considers the basic drives of man...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 of the other drives. For every drive wants to be master—and it attempts to philosophize in that spirit.” (BGE 7, 203-204). This idea, where knowledge itself is but a reflection of an unknowable drive is expressed even more forcefully in *On The Genealogy of Morals (GM Preface, 2)*: “A fundamental will to knowledge is the root in which... our ideas, our values, our yeas and nays, our ifs and buts, grow out of us with the necessity with which a tree bears fruit—related and each with an affinity to each, and evidence of one will, one health, one soil, one sun.”<sup>8</sup> If every idea and every value grows out of a “soil” or “sun” that cannot, in principle, be fully exposed, then any attempt to reveal the conceptual, empirical or psychological constituents of that same soil or sun would give rise to an endless investigation—what we take are as fundamental constituents of our drives are in fact manifestations of the drives themselves.<sup>9</sup>

If, as Nietzsche confirms elsewhere, philosophy is simply a confession of the drive constitution of its author, how do we objectively confirm the drives responsible for said philosophizing?<sup>10</sup> The clear answer is we can't. Drives or something like them must be postulated to make sense of our psychology. I chose to expound on three elements of our collective human psychology here. First, drives are transformed animal instincts, so some aspects of drives will never be fully articulable for human beings. The remaining two aspects are 2) drives run along fixed routes with definite aims and 3) they urge agents to take these routes independently of deliberation. As I will argue in the last section, the best way to

---

<sup>8</sup> Friedrich Nietzsche, *On/Towards A Genealogy of Morals*, GM preface 2, 452.

<sup>9</sup> See also GM III 7: “Every animal therefore la bete philosophe too—instinctively strives for an optimum of favorable conditions under which it can expend all its strength and achieve it's maximal feeling of power; every animal abhors, just as instinctively and with the subtlety of discernment that is ‘higher than all reason’, every kind of intrusion or hindrance that obstructs or could obstruct this path to the optimum.” (GM III 7) Similar ideas regarding the fundamental inability to understand the fundamental assumptions that make knowledge possible can be found in the middle works as well. In *Gay Science*, section 111, for example, Nietzsche writes: “He who did not know how to discover the identical sufficiently often in regard to food or to animals hostile to him, he who was too slow to subsume, too cautious in subsuming, had a smaller probability of survival than he who in every case of similarity at once conjectured identity.” For an extensive treatment regarding the reflexivity of Nietzschean drive theory, see my book, Brian Lightbody, *Nietzsche's Will to Power Naturalized: Translating the Human into Nature and Nature into the Human*. (Lanham Maryland: Rowman and Littlefield, Lexington Books), 2017,

<sup>10</sup> See GM II 12 and BGE 36.



confirm the existence of drives is to think of them as givens: we have good, albeit indirect, evidence to think drives exist. Before looking at this answer in greater detail, I wish to examine Stern's criticism of Nietzsche on this front, as it will help to situate the position more clearly and provide a bridge between Nietzsche's and Socrates's respective psychological models.

## II Stern's criticism of Nietzsche.

In his article, Stern explores the implications of the animality of drives and their relationship to consciousness. He discovers that even this seemingly non-contentious claim that drives have an animal base, are goal-oriented, and manifest an urgency for expression are insufficient to replace consciousness. He writes:

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 non-human animal, are not the sorts of things that we'd be inclined to say are naturally seeking expression...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. (Stern 2015, 127)

To take but one example that buttresses Stern's case but one that he does not treat in full, consider the instinct for "luxury." In *On the Genealogy of Morals* III 8, Nietzsche claims that inventive spirits throughout the ages have had to resist natural instincts towards "a love of luxury and refinement or an excessive liberality of heart and hand" (Nietzsche, 2000, GM III 8). Moreover, these instincts were checked because of a more dominating spiritualizing instinct: "The three great slogans of the ascetic ideal are familiar: poverty, humility, chastity. Now take a close look at the lives of all the great inventive spirits: you will always encounter all three to a lesser degree" (Nietzsche, 2000, GM III: 8). But as Nietzsche explains, these dispositions towards the finer things in life are negated not through the imposition and implementation of virtue, whether construed as Greek *Arete* or Christian moral practices or even in a more ordinary sense like forming a resolution (e.g. I shall not succumb to my bodily appetites and will remain chaste ) but rather because of a "dominating spirit." "It was the dominating spirit whose demands prevailed against those of all the other instincts—it continues to do it; if it did not do it, it would not dominate. There is thus nothing of virtue in this" (Nietzsche, 2000, GM III: 8). Notice that the character of a dominating spirit has demands. But these demands are not conscious goals deemed more important than one's primal instincts; rather, they

are just another kind of instinct, another kind of drive. The difficulty here is in coming to terms with the biological basis for the drive to pursue, let's call it, the finer things in life.<sup>11</sup>

Let's be clear about the problem. The problem is not that persons may be drawn to pursue such opulent things unconsciously—Marcuse in *Eros and Civilization* for one provides an illuminating account of how a subject's libidinal desires may be harnessed and redirected towards some commercial product (Marcuse, 1966). The point I return to is the one made by Richardson above. How is one expected to view the drive for luxury through an animal framework where, from this creaturely perspective, each animal has defined instincts that unconsciously impel and direct it to some object? Now Richardson may claim that the drive of luxury is a transformation and refinement of a plastic primeval drive for acquisitiveness—magpies, it is often, said, instinctively seek out shiny baubles—but if this is indeed the ancestral instinct that platforms the drive in question, it is not clear how the evident and incredible recalibrating of this once Ur-drive came to be. Echoing a *leitmotif* of Stern's: Where do drives, as the explanatory tokens of human psychology, (as quoted above by Katsafanas) come in to play to explain behavior?

According to Stern, thinking about the drive to luxury as a transformation of some animal instinct leads to another issue: what is the relationship between deliberation and drives? Clearly evaluation, reflection, and analysis are needed to understand what object or service meets the criterion of extravagance and/or pampering per GM III 8's drive to luxury, over some other. In addition, how one attains such items also requires careful consideration less one's goal of obtaining a luxurious life instead becomes a life of pain and misery. Such musings, however, require conscious reflection. But then, what is the relation between drives and consciousness? And was it, not the case that drives were introduced to offer an *explanans* for the hardest of all *explananda*, namely conscious itself?

On this problem Stern notes:

Drives, recall, are the things that propel animals to behave quasi-rationally, quasi-purposively and quasi-expediently. But the salient feature of the animal, unlike the human, is that the non-'quasi'-versions are lacking: the human building a home, unlike the bee, does, *prima facie*, have consciousness, reason, instruction, a plan, prior models and so on. Given *prima facie* consciousness et al., 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. (Stern 2015, 126)

Stern presses his case against Nietzsche by bringing in expert testimony from none other than Socrates himself. Socrates' testament is particularly pertinent because, as Stern reminds the reader, Nietzsche's

<sup>11</sup> For an extended discussion, See Brian Lightbody, *A Genealogical Analysis of Nietzschean Drive Theory* (London: Palgrave Macmillan), 2023.

intended purpose of drive theory is to challenge the Socratic picture of moral psychology. Stern spells this out in the following, where he is discussing Socrates' predicament after his trial, as described in the *Apology*. The discussion summarizes Socrates's deliberation in *Phaedo*, where Athens' gadfly contemplates the morality of escaping from his jail cell. On Socrates' account of the relationship between what he thought and what he did, the following seems to be true or so holds Stern:

i) Socrates could have chosen to do different things; ii) Socrates correctly weighed up his options and iii) acted according to which option he thought was right; iv) having acted, Socrates' motivations are plain for him to see and to be made available to others such that v) he can be judged morally based on what he chose to do. (Stern 2015, 122)

Although Stern does not quote the passage in question, it is instructive to spend some space outlining Socrates' moral psychology. As Socrates mentions in the *Phaedo*, the actual causes for his being in the jail cell awaiting his fate instead of fleeing to some other city-state, such as Thessaly, are not his flesh, bones, and sinews, that is, his body. If someone were to say that Socrates remains in the jail cell because of where his body is, we would think the person is mad or that he hasn't understood the question. Socrates himself explains the real reason he has refused his friends' suggestion to escape: "The real causes, which are, that the Athenians decided that it was best to condemn me, and therefore I have decided that it was best for me to sit here and that it is right for me to stay and undergo whatever penalty they order." (Plato, *Phaedo* 99a) Notice that the explanation is succinct and perfectly lucid. It comprises a host of elements united into a logical and rational explanation that forms a narrative we, readers, can readily understand. In comprehending the reasons for Socrates's decision, we enter into what Wilfrid Sellars calls the space of reasons— a normative sphere where epistemic claims are endorsed or sanctioned by our best judgment and where our best judgments cannot be further reduced to crude naturalistic causes.<sup>12</sup> Let's call this claim the irreducibility thesis.

IR: Every normative decision  $x$  is initiated by reasons that are not solely reducible to naturalistic causes.

Proof for the irreducibility thesis is provided by simply reflecting on whether Socrates's action to stay is justified or not. As readers of the *Phaedo*, we can understand and justify (or condemn) Socrates's decision because we can reason along with him and compare the weights Socrates assigns to the pros and cons of his alternatives and then evaluate them to what we think the right thing to do might be.

Yet Stern reminds us that Nietzsche's drive theory refutes the irreducibility claim. Drive theory reveals an even more thorough and illuminating explanation regarding the *true causes* of Socrates's staying put. But if this is what Nietzsche is doing then surely it is incumbent on him (and his defenders to explain), "(a) what sorts of things 'drives' are; (b) how many, and how much, we know about them; and

<sup>12</sup> See Wilfrid Sellars, *Science, Perception and Reality* (London: Routledge and Kegan Paul, 1963), 169

(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 choices?" (Stern, 2015, 123) Nietzsche, however, does not answer any of these questions, and thus, Nietzsche's model is in no way preferable to what appears on the surface to be an intuitive, parsimonious, common-sense model of moral psychology where deliberation is causally efficacious full stop.<sup>13</sup>

At this juncture, Stern, playing the role of prosecutor, rests. However, before we condemn Nietzsche's drive model to the gallows of either incoherency or vacuity, it is vital to return to the question both Socrates and Stern pose: What is the real cause for Socrates sitting in this jail cell? Why did Socrates refuse his friends' offer to escape? It would not be deliberation *simpliciter* as Stern thinks it is. The answer is far more complex.

In order to think about Socrates's refusal to accept his friends' help in escaping his fate, we need to examine the Gadfly's moral psychology in greater detail, particularly his daimon, which serves as a lynchpin between the Forms and deliberation in much the same way as the phenomenological craving that manifests in subjects when a drive takes hold of them serves as the bridge between the drive's biological platforming and the channels for its expression. In what follows, I suggest there are more similarities than differences between the respective philosophical psychologies of Socrates and Nietzsche than a first glance would seem to suggest. I demonstrate that the primary currency of each (e.g., Socrates's Forms and Nietzsche's drives) are advanced as "givens"—axiomatic statements that are meant to cohere with what each philosopher takes to be other well-supported claims. Let me explain.

### Section III: Socrates's Moral Psychology: Daimon and Elenchus

Socrates's Daimon is mentioned in many places in Plato's dialogues. (See *Phaedrus* 242 b; *Apology* 31d; 40a; *Euthyphro* 3b, *Symposium* 175b, 202e-203a and *Hippias Major* 304 b-c). What precisely this *daimon* is has been the subject of fierce debate in the secondary literature.<sup>14</sup> Perhaps the most extensive discussion

---

<sup>13</sup> Stern argues: "Now Nietzsche does provide answers to this question, but he is delinquent when it comes to delivering a consistent solution. Again, quoting Stern, "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." Stern, "Against Nietzsche's Theory of Drives", 123.

<sup>14</sup>Kraut, R. 1981. "Plato's *Apology* and *Crito*: Two Recent Studies" *Ethics*, 91: 651-664. In Martha Nussbaum's "Introducing a New God: Socrates and His Daimonion Commentary on Edmunds. Proceedings of the Boston Area Colloquium of Ancient Philosophy 1: 231-240, 1985, Nussbaum argues for an ironic reading of these passages. She writes "that an reference to the daimonion is but an ironic way of alluding to the supreme authority of dissuasive reasoning and

of the powers and function of this *daimon* is provided in *the Apology*. After Socrates' has been condemned to death, he explains to the jurors a strange phenomenological experience:

A surprising thing has happened to me jurymen you I would rightly call jurymen. At all previous times my familiar prophetic power, my spiritual manifestation, frequently opposed me even in small matters when I was about to do something wrong but now that as you can see for yourselves I was faced with what one might think and what is generally thought to be the worst of evils, my divine sign has not opposed me either when I left home at dawn, or when I came into court, or at any time that I was about to say something during my speech. Yet in other talks it often held me back in the middle of my speaking, but now it has opposed no word or deed of mine. What do I think is the reason for this? I will tell you. What has happened to me may very well be a good thing, and those of us who believe death to be an evil are certainly mistaken. I have convincing proof of this for it is impossible that my familiar sign did not oppose me if I was not about to do what was right." (*Apology*, 40a-c)

If we notice in Socrates's speech, the *daimon* has causal powers. First, it can stop Socrates in his cognitive tracks: as Socrates remarks, the *daimon* would sometimes appear to him just before he undertook some small deed or action. As Gerd van Riel puts it, "The daimōn only utters a 'Stop!' without explanation. This inarticulate message leaves Socrates in ignorance: He wanted to do something with the best of intentions, but the daimōn held him back, purely by virtue of its presence."<sup>15</sup> For Van Riel, Socrates's philosophical psychology is divided into the daimonion and elenchus. The daimonion has province over Socrates's actions but "...it never has anything to do with the convictions or opinions of Socrates; it acts exclusively on the actions envisaged by Socrates, and even more specifically, on actions that are not good. Socrates's dialectic of Enlenchos has domain over logic, deliberation or the regime of truth."<sup>16</sup>

Yet it is evident from Socrates's speech above that the so-called clear demarcation between action and deliberation Riel describes cannot be correct. While it may be true that the *daimon* never has anything to do with the opinions or convictions of Socrates *stricto sensu* it clearly has the power to interrupt Socrates in the middle of a speech, preventing him from reaching conclusions the *daimon* feels are unwarranted. That capacity entails that the *daimon* has evaluative powers after all because it enters into Socrates's "space of reasons" as noted above, and somehow dissuades the gadfly of Athens from proceeding down a course of inquiry he initially thought was correct. But how is it able to do this?

---

elenctic argument." (234). McPherran counters by arguing "...to reread them all as sly tongue-in-cheek verbal pandering or as simply allegorical in intent would be to employ a principle of interpretation that, once loosed upon the texts, would no know end, rendering every Socratic utterance fatally indeterminate." McPherran, M. *The Religion of Socrates*. University Park: The Pennsylvania State University Press, 1996, 7.

<sup>15</sup> Gerd Van Riel "Socrates' Daemon: Internalisation of the Divine and Knowledge of the Self", *Apeiron: A Journal for Ancient Philosophy and Science*, June 2005, Vol. 38, No. 2 31-42, 38.

<sup>16</sup> Van Riel, 38.

To answer this question regarding the daimon's strange evaluative powers, we must turn to the *Phaedo* once again and, in particular, to Socrates's response (100a-e) to Cebes's question about why he refused his friends' offer to escape to complete the initial answer given in 99—a critical section Stern leaves untreated. There, Socrates says, "I assume the existence of a Beautiful, itself by itself, of a Good and a Great and all the rest." (*Phaedo* 100b) Parsing this in Plato's middle-period metaphysics, the Forms represent perfect cognitive paradigms we come to remember (*Amanesis*) here on earth. We go through a cycle of birth, death, and rebirth, and when our soul is separated from our bodies, we see the Forms as perfect ideas in the heavens (*Phaedo*, 72e-77e) Without getting into unnecessary detail about Plato's middle theory in works like *Republic* and *Symposium*, here in the earlier middle work, the *Phaedo*, Socrates remarks that the idea of Equality cannot be gathered by comparing two seemingly identical things no matter how seemingly exact and of the same size they appear to our senses for we will always recognize that no two things are ever truly the same. To illustrate the truth of his position, he uses the example of two sticks that appear to be the same size at first glance but are unequal in length upon measuring them. The Forms allow us to know why the sticks are equal or at least relatively so, because the Form of Equality allows us to make measurements *simpliciter* (*Phaedo*, 74b-c). True measurement would be impossible without some standard that is the same in itself. That line of reasoning applies to standards of measurement, like metre sticks, which is why the *metre des Archives*, the first physical template of the standard meter, was copied to make meter sticks ever since. The same also holds true for perfect ideas. As Socrates concludes: "Consider then, he said, whether you share my opinion as to what follows for I think that, if there is anything beautiful besides the Beautiful itself, it is beautiful for no other reason than that it shares in that Beautiful, and I say so with everything. Do you agree this sort of cause? I do." (*Phaedo* 100c).

The abductive nature of what we might call the crux of what scholars have called the recollection argument 72e-77e, as well as other parts of the dialogue, have received extensive treatment in the secondary literature<sup>17</sup> Jetli, for example, has provided a penetrating and illuminating analysis of Socrates's

---

<sup>17</sup> Jetli, P. "Deduction–Abduction–Induction Chains in Plato's *Phaedo* and *Parmenides*", In: Magnani, L. (eds) *Handbook of Abductive Cognition*. Springer, Cham. 2023. See also, ARAÚJO, H. F. de; MONTENEGRO, M. A. de P. Perception and Memory in Plato. *Kalagatos*, [S. l.], v. 21, n. 2, p. eK24045, 2024. Disponível em: <https://revistas.uece.br/index.php/kalagatos/article/view/13376>. Acesso em: 13 jul. 2024. They write: "As we shall see, the philosophy of Plato, centered on the hypothesis of the existence of an Intelligible reality (Form), is based on the need to explain one's own sense experience, in all its nuances, whether in the ontological field, in which the Forms are the cause (*aitia*) of the sensible, or in the epistemological field, where the philosopher undertakes the task of defending that sensible perception is effective because there are notions in the individual that are prior to birth (i.e., stored in the soul), capable of organizing the data captured by the senses. (p. 2)

argumentative cycle, as he calls it, of deductive to abductive followed by inductive chains of inference, *et repetens*, throughout the *Phaedo*. Indeed, this cycle is demonstrable near the very beginning of the dialogue, where Socrates attempts to demonstrate there must be a different substance, other than the body that grounds our subjectivity from the fact that we experience two very different drives within us: a desire for truth on the one hand and a desire for all bodily things. (*Phaedo*, 64c-65d). The assumption that Socrates trades on (as Nietzsche well understood) is that oppositional desires entail two contrary things: desires cannot exist without being attached to the thing doing the desiring. If that makes sense, then although we as subjects seem to be composed of contrary things, we may be able to separate them, and it is this belief that because there are conflicting values, there must, subsequently, be conflicting things. As Nietzsche clarifies: “The fundamental faith of the metaphysicians is the faith in opposite values.” (BGE 2)

Nietzsche's diagnosis is not so much incorrect as inexact, though, for Socrates fully realizes that such reasoning (opposite values imply opposite things) is not conclusive; it is an abductive hypothesis that merely warrants further exploration. He notes: “There is likely to be something such as a path to guide us out of our confusion, because as long as we have a body and our soul is fused with such an evil, we shall never adequately attain what we desire, which we affirm to be truth.” (*Phaedo*, 66b-c) The clearest articulation of what some scholars have called the conflict argument in Plato's work is probably found in the *Republic*, 436e.<sup>18</sup> Socrates says: “No such statement will disturb us, then, or make us believe that the same thing can be, do, or undergo opposites, at the same time, in the same respect, and in relation to the same thing.” (*Republic* 436e 1068).

However, perhaps the work's most straightforward and thoroughgoing and yet compact display of abduction in the *Phaedo* or perhaps even in Plato's oeuvre is Socrates's eclipse analogy, which starts at 99C. It is vital to explore this argument in detail as it greatly clarifies Socrates's abductive thinking. Indeed, Pamela Huby in her article, “Phaedo 99d-102A” argues that the passage is an instance of Simmias's early call to “...studying things in *logoi* and, in particular, selecting on each occasion the logos that seemed to be the most reliable.”<sup>19</sup> Socrates not only hears but answers Simmias's call. In the passage, Cebes is clearly unsure what Socrates is talking about in discussing first causes and the like (e.g. the Beautiful in itself, Equality and so on) as we saw on 100c. And so the gadfly offers the following helpful explanation:

---

<sup>18</sup> See R.E. Stalley, “Plato's Argument for the Division of the Reasoning and Appetitive Elements within the Soul”, *Phronesis* 5 20.2 (1975) 110-128 and Tim Robison, “Plato's Separation of Reason and Desire,” *Phronesis* 16 (1971) 38-48. See my book, Brian Lightbody *Dispersing the Clouds of Temptation*, (Eugene Oregon 2015) pp. 90-109.

<sup>19</sup> Phaedo" 99 D-102 A Pamela M. Huby *Phronesis*, Vol. 4, No. 1 (1959), pp. 12-14

I would gladly become the disciple of any man who taught the workings of that kind of cause. However, since I was deprived and could neither discover it myself nor learn it from another, do you wish me to give you an explanation of how, as a second best, I busied myself with the search for the cause, Cebes? I wish it above all else he said. After this, he said, when I had wearied of investigating things, I thought that I must be careful to avoid the experience of those who watch an eclipse of the sun, for some of them ruin their eyes unless they watch its reflection in water or some such material. A similar thought crossed my mind, and I feared that my soul be altogether blinded if I looked at things with my eyes and tried to grasp them with each of my senses. So I thought I must take refuge in discussions and investigate the truth of things by means of words.” (99 C *Phaedo*)

In the words once more of Huby, “It has usually been supposed that what they do not understand is his method, but his words can equally mean that they do not see what the particular *logos* (here understood as theory or provisional ground) is which he has *chosen* and goes on to explain—the Theory of Forms. Taking this as his *logos*, he will now consider its consequences (e.g., proving the soul's immortality). This shows the wide utility of the *hypothesis*.”<sup>20</sup> There are two essential points to notice in Huby's quote: 1) the rational principle is a hypothesis, and 2) it is chosen seemingly from other possible principles. However, I want to identify another notion here: why the principle is a freely chosen hypothesis ( an axiomatic given that Socrates asks his friends to accept) and why it can be corroborated, to some, degree but never directly verified. To answer this question, we must understand the analogy.

It is critical to draw out the analogy's salient features to comprehend it. There are three features I wish to elucidate. First we have the eclipse of the sun. Eclipses damage the eyes because the natural physiological mechanism to squint and reduce overexposure to light waves is significantly mitigated. One is tricked into thinking that because the moon appears to block all the sun's rays, one is not exposed to harmful UV light, which, after prolonged exposure, causes one to become blind. To simplify, I call this the blindness claim.

Of course, Socrates would not have been aware of the scientific explanation for eclipses noted above, but he was aware that the sun, as the ultimate cause for sight itself, could damage our ocular capacities and render them useless. This leads us to the second feature of the analogy: the sun is the ultimate cause of our collective capacity to see. More must be said about describing the sun as the fundamental cause of sight because analogously, Socrates will trade on this fact to discuss the cause behind our ideas, such as the Beautiful in itself, justice, etc. I call this second feature of the sun the first cause feature.

I now turn to the third feature. Eclipses offer us a rare chance to look at various aspects of the sun, which we would not otherwise have the opportunity to see because of our natural inclination to squint. For example, typically, when we look directly at the sun, we only see the photosphere, but with

---

<sup>20</sup> Huby, 13.



an eclipse, we see both the sun's corona and the coronosphere. If we wish to look at these features without becoming blind, then as Socrates's suggests, we need some reflective source, such as a pool of water or some other material. Such a reflective resource allows us to see the features of a particular eclipse and compare these aspects to future, indirectly observed eclipses. I call this aspect the reflective feature.

With these three physical features of the eclipse in mind, I now examine how each may be transposed into a symbolic key. I look at the blindness claim first. The blindness claim asserts that individuals who stare at an eclipse for an extended period may so damage their eyes that they go blind. Socrates uses this imagery to explain to Cebes his struggle to find the first causes of things. He exclaims "when I had wearied of investigating things, I thought that I must be careful to avoid the experience of those who watch an eclipse of the sun, for some of them ruin their eyes." The analogical line of reasoning that follows seems to be this: just as one may become blind by staring at the sun for extended periods of time, one may lose one's capacity to form correct philosophical judgments by over-analyzing ideas without being able to track if one is making progress in coming to terms with the idea in question. Consider what it would mean to run a long-distance race without knowing how many miles constituted the challenge. Such a runner would surely grow weary without knowing the total miles to run or how many miles have already been covered. This example is analogous to the problem Socrates mentions here: if I am looking for the cause of an idea such as beauty, courage, or piety but cannot know if I am making progress towards finding the authentic source behind such notions—because I am conceptually blind—then I will no doubt become exhausted from such a strenuous mental investigation.

In order to prevent such intellectual weariness, there must be a standard to track my progress. The second claim, the first cause feature, functions as that standard; it may now be connected to the abductive argument Socrates made about equality earlier in 74 b-c. As noted above, Socrates's argument postulates a primordial idea of Equality, which is equal to itself. This postulation holds that when we see two seemingly equal things, we are reminded of the Form of Equality itself. Since we clearly do have the idea of Equality, and yet no two things are perfectly equal, it must be the case that the idea is not of this realm—we arrived at it before we were born. The first cause feature of the sun analogy generalizes this line of reasoning: just as the sun is the first cause of all sight, these ideas or Forms are the cause for all concepts. Only by postulating the existence of such ideas as the Beautiful in itself can there be progress.

I now turn to the last feature: the reflective claim. This claim is intimately connected to the discussion of Socrates's *daimon*, so it is vital to discuss them in tandem. The reflective claim suggests that during an eclipse, one can see features of the sun that are not ordinarily viewable. However, viewing these features using reflective material is best, as it will lessen the likelihood of one becoming blind. In translating this component of the analogy symbolically, the reflective pools one uses to view an eclipse

are the words one may use to gain a better purchase on the absolute principle (Form) one is investigating. As Socrates states, "So I thought I must take refuge in discussions and investigate the truth of things by means of words." In other words, philosophical investigations regarding the nature of virtues like courage (*Laches*) or piety (*Euthyphro*) are dynamic processes. Socrates developed a unique dialogical process to investigate such essences: elenchus. The goal of an elenctic inquiry is to discover the essence of a virtue, X, by formulating that essence into a definition that would apply to all examples of X. As any reader of the early Platonic dialogues knows, such a request is nearly impossible to fulfill, and all of Socrates's interlocutors fail badly in attempting to satisfy his demand.<sup>21</sup>

However, the judgment as to whether the words get the matter at hand right is not something that deliberation itself can make. Following the eclipse analogy, philosophical musings must reflect the principles that give them movement or kinesis. How does Socrates know if the philosophical discussion is not developing as it should be if he cannot directly check the corresponding Form with the words that reflect it via the eclipse/sun analogy? Here, the *daimon* comes into play by arresting Socrates's speech and forcing him to reconsider his thinking insofar as it mirrors the appropriate Form it reflects. The *daimon*, which has access to the Forms, makes the judgment non-verbally and intuitively: a definition is on the right pathway provided that the *daimon* does not say Stop!

If that's right, then elenchus is necessarily tied to questions about fundamental principles that it cannot directly interrogate or even doubt. We can question the features and aspects of what we see being reflected as it were, as the discussion brings these features to light, but we can not doubt the very thing that is the cause of the reflection itself. To do so would be wildly incoherent: How could one uphold that the sun is being reflected in a pool and yet believe that no such thing as the sun is being so reflected?

Yet, this is exactly what some do when we cash out the analogy metaphorically. To wit, there are interlocutors Socrates's engages with, such as Alcibiades, Gorgias, Callicles, and Thrasymachus, who question the very existence of the virtue *their discussion merely reflects!*<sup>22</sup> Socrates argues that not only is such

---

<sup>21</sup> As Richard Robinson demonstrates the reason why such a task is destined to fail is because Socrates does not differentiate between two senses of interpreting a What is X question: (1) On the one hand, many passages suggest that all he [Socrates] wants is a mark that shall serve as a pattern by which to judge of any given thing whether it is an X or not. In the *Euthyphro* (6 E) he describes his aim in just this way. (2) In many other passages, however, Socrates' purpose in asking What-is-X? is evidently not, or not merely, to distinguish X from everything else. It is to get at what he calls the essence or form of X ..."

Richard Robinson, *Plato's Earlier Dialectic* (Ithaca, New York: Cornell University Press, 1941). 56- 57.

<sup>22</sup> I cannot discuss all these thinkers in detail but consider what Callicles states regarding the relationship between bodily desires and reason in 491e-492a of *Gorgias*: "When they are as large as possible (bodily appetites), he ought to be competent

a manner of argumentation wrongheaded and causes one to grow spiritually weary—as noted because there is no way to track intellectual progress—but also has disastrous moral consequences in that one may become so conceptually blinded that one espouses and practices beliefs that are harmful to the soul itself.<sup>23</sup> Plato demonstrates that each of the above interlocutors becomes corrupted because they question the existence of timeless, eternal virtues that cannot be understood via deliberation alone. For Plato, we can make philosophical and moral progress only by postulating the existence of eternal, permanent Forms beyond this world of flux and change. In the next section, I will discuss the aspects of Socrates's and Nietzsche's respective abductive postulates. Where each philosopher seems to offer different models of mind, at least at first glance, I will show they have much in common.

#### Section IV: Abduction

Abductive arguments are a form of non-deductive reasoning. With deductive reasoning, a conclusion necessarily follows from premises provided those premises are couched in a valid argument; the conclusion, in other words, is 'contained' in the premises. Abduction, in contrast, is similar to induction in that it is ampliative: it is a rational process that leads to a conclusion that is not already enclosed in the original premises.<sup>24</sup> However, unlike induction, which moves from an observed common trait about a sample size of things (E.G., every swan I have seen has been white) to a general rule about those very things observed (therefore, all swans are white), abductive reasoning is indirect. We come to a particular conclusion to explain some phenomenon not through direct observation (as with the swan example above) but indirectly. In other words, I think: 'Given I do not have direct confirming or

---

to devote himself to them by virtue of his courage and intelligence, and to fill him with whatever he may have an appetite for at the time." (Gorgias 491e-492a trans. Zeyl). Callicles is so beyond the moral and rational pale that even Socrates's dialectical method is ineffectual or so I have argued. See my paper Brian Lightbody, "On Becoming Fearful Quickly: A Reinterpretation of Aristotle's Somatic Model of Socratean Akrasia." *The Journal of Ancient Philosophy*, Vol. 17 Issue 2, 2023, 134-161 for further analysis.

<sup>23</sup> To give but one example, consider Callicles's advocacy of Pleonexia, the position that advances the thesis that one should allow one's appetites to get as large as possible and not restrain them. Moreover, Callicles notes that, "When they are as large as possible, he ought to be competent to devote himself to them by virtue of his courage and intelligence, and to fill him with whatever he may have an appetite for at the time." (Gorgias 491e-492a trans. Zeyl). Even the individual with little or no philosophical training would clearly see that such an ethos would have disastrous results for the person who ascribes to such a view.

<sup>24</sup> See my book Brian Lightbody, *The Problem of Naturalism: Analytic Perspectives, Continental Virtues*. (Lanham, Maryland: Lexington Books, a division of Rowman and Littlefield), 2013, 34.

disconfirming evidence for the phenomenon, what would be the best explanation for the occurrence of the phenomenon based on the current set of beliefs I select to be relevant in this case?' As Charles Peirce summarizes this way of reasoning, "abduction is the process of forming explanatory hypotheses. It is the only logical operation which introduces any new idea."<sup>25</sup> Simplifying, abduction is a form of reasoning that makes an inference to the best explanation.

A concrete example will help illustrate the point. If I come across a car parked with its trunk open, I think the driver either 1) intentionally opened the trunk but then forgot to close it, 2) left the trunk open intentionally, or 3) some thief intentionally broke the lock on the trunk. I do not think the trunk randomly popped open "on its own." Though I do not have direct evidence to confirm either of the three possible causes above, based on my background information about cars and people's intentions, one of the above three possibilities is most likely the reason for the trunk being open. Abduction is ultimately intimately connected to providing the best explanation based on a set of reasons one takes to be given; they are reasons I will not challenge unless I have direct disconfirming evidence.

In examining the abductive explanations to account for Socrates's model, I will demonstrate that it relies on questionable "givens." One of the arguments Socrates provides for the existence of the Forms is the argument for absolute equality. When we compare two sticks that appear to be equal, we can form this judgment only because we possess, via *Amanesis*, the knowledge of the Form of Equality. In order for this explanation to work, however, it presupposes, at a minimum, several metaphysical and epistemological assumptions, such as the existence of timeless Forms, non-empirical access to a very special kind of knowledge, and metaphysical dualism: the soul and body are two distinct things. Socrates's eclipse analogy further demonstrates why the Forms are not directly observable. They are primal causes that platform, in Stern's nomenclature, "deliberation" *per se* and, more specifically, Socrates's elenchus method, designed to root out the essence of a specific virtue. Without the postulation of an idea that is timeless and identical to itself, moral progress is impossible, weariness sets in, and the temptation to follow a pathway that corrupts the soul becomes likely, or so Socrates reasons.

Nietzsche's abductive processes are more sophisticated and substantive than those of the ancient Athenian philosopher; they, too, are givens but ones that are naturalistically inflected. They are speculations about early human anthropology that lack direct confirming evidence but cohere with a scientific worldview of human evolution. I cannot delineate all the givens here, however, but they are nicely captured by the German philosopher on the first line of GM II 1: "To breed an animal with the

---

<sup>25</sup> Douven, Igor, "Abduction", The Stanford Encyclopedia of Philosophy (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2021/entries/abduction/>>.

right to make promises—is not this the paradoxical task that nature has set itself in the case of man?” As Nietzsche unfolds his philosophical anthropology in the *Genealogy*, he postulates that drives are transformed animal instincts—they are the product of various somatic technologies such as laws, walls and punishment for breaking said laws, and softer social norms and regulations. The givens in Nietzsche's genealogical account are these aggressive animal instincts that direct the animal body to track and pursue various targets (e.g., there is a hunting instinct, an adventure instinct, a violence instinct, a sex instinct, etc. GM II 16).<sup>26</sup> Yet another given is the capacity for these same instincts to turn inward when blocked from flowing in their appropriate channels. These instincts, when obstructed, transform into drives that take on different external targets but, more importantly, serve to carve out the soul.<sup>27</sup> The soul, Nietzsche reminds us, before the advent of civilization, was as thin as “...if stretched between two membranes.”<sup>28</sup> These givens perfectly fit the scientific outlook of Nietzsche's day, and indeed, the question that opens GM II 1 resonates with readers today.

In comparing the two models, Socrates's is much like the first explanation to account for the car with its trunk open. Some might think this explanation is the best because only a driver or another person can open or close a trunk. A trunk's opening or closing is predicated on a remembered or forgotten intention (remember to close the trunk!). Socrates's account for the existence of all physical things here on earth is to postulate an abstract non-material entity that guides him to arrive at the right conclusion via the powers of his mysterious *daimon* and *elenchus*. Although he has no control over when the *daimon* appears, he does have control over the development of his reasoning. He also has control over initiating

---

<sup>26</sup> Lightbody, Brian. “The Passive Body and States of Nature: An Examination of the Methodological Role State of Nature Theory Plays in Williams and Nietzsche.” Special Issue: Philosophical Genealogy from Nietzsche to Williams, *Genealogy* 5 (2), 38, 2021, 1-15.

<sup>27</sup> This process of “soul carving” that Nietzsche discusses in section GM: II, 16 is often referred to as the Internalization Hypothesis. For two different readings of this Hypothesis, see my article Lightbody, Brian. “Twilight of the Genealogy or A Genealogy of Twilight? Saving Nietzsche's Internalization Hypothesis from Naïve Determinism.” *Philosophical Readings*, Volume 3, Sept 2021 183-194 (2021b). This article attempts to justify a literal reading of internalization. The second interpretation is an implexic reading of internalization. See Brian Lightbody, Chapter Four The Internalization Hypothesis: A New Reading, in *A Genealogical Analysis of Nietzschean Drive Theory* (London: U.K. Palgrave Macmillan), 2023.

<sup>28</sup> Nietzsche, GM: II 16: “The entire inner world, originally as thin as if it were stretched between two membranes, expanded and extended itself, acquired depth, breadth, and height in the same measure as outward discharge was inhibited.”

an action, such as when he decides to stay in his jail cell. On that score, Stern is correct: intentions platformed by deliberation under the Socratic model are causally efficacious.

However, such causal efficacy is not attached to the body, at least not meaningfully. For Nietzsche, intentions are the product of an unconscious drive's aim. Moreover, there is reason to believe that drives might be mapped onto a contemporary neuroscientific understanding of the brain. Two prominent scholars in the literature have argued that J. Panksepp's Affective Neuroscientific Model aligns with Nietzsche's theory of drives.<sup>29</sup> Panksepp identified seven primary emotional command systems in humans and mammals: SEEKING/expectancy, RAGE/anger, FEAR/anxiety, LUST, CARE/nurturing, PANIC/sadness, and PLAY/social joy.<sup>30</sup> While I cannot delve into Panksepp's model in detail here, these systems aim to ground the emotional nature of mammalian behavior and the corresponding assimilative, inquisitive, and appropriative actions seen in animals on a neurophysiological basis.<sup>31</sup> More specifically, these systems are considered primary because they are rooted in the subcortex. One of Panksepp's experiments suggests this, as rats that had their cortex surgically removed still exhibited some of these basic emotions.<sup>32</sup>

One point is critical to note on this score. Panksepp's basic affective systems cannot be understood without the drives that serve as their representational correlates. As Nietzsche contends, the drive for knowledge underpins the cognitive craving to make neuroscientific discoveries. Drives serve the same purpose as Socrates's Forms: they are the causes of our behavior that we cannot observe directly. They platform our representational capacities for scientific framing *per se* because without an unconscious

---

<sup>29</sup> Mattia Riccardi and Rex Welshon have utilized Panksepp's understanding of these systems to provide neuroscientific support for Nietzsche's drive theory. Welshon, for instance, argues that these systems are, "...entirely congruent with Nietzsche's claims about the causal efficacy of drives." See Rex Welshon, *Nietzsche's Dynamic Metapsychology: This Uncanny Animal*, (London: Palgrave Macmillan, 2014), 121. While Riccardi states that, "The parallels between this picture (Panksepp's) and Nietzsche's drives -cum-affects psychology are obvious enough." Riccardi, *Nietzsche's Philosophical Psychology*, 68. For an alternative reading of Panksepp in relation to Nietzsche, see Brian Lightbody, Chapter Three, "Drives in the Secondary Literature" *A Genealogical Analysis of Nietzschean Drive Theory*, 103-107.

<sup>30</sup> Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York, NY: Oxford University Press.

<sup>31</sup> Panksepp, J. (2011). Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. *PLoS ONE* 6:e21236. doi: 10.1371/journal.pone.0021236

<sup>32</sup> Panksepp, J., Normansell, L., Cox, J., and Siviy, S. (1994). "Effects of neonatal decortication on the social play of juvenile rats." *Physiol. Behav.* 56, 429-443. doi: 10.1016/0031-9384(94)90285-2

force propelling us to investigate the nature of things, so to speak, reason, all by itself is inert. Only a drive, Nietzsche contends, provides the impulsive thirst for knowledge: "To be sure sure: among scholars who are really scientific men, things may be different—"better", if you like—there you may really find something like a drive for knowledge, some small, independent clockwork that, once well round, works on vigorously without any essential participation from all the other drives of the scholar." (BGE 6) Thus, even if Panksepp is correct, we cannot understand these systems without utilizing the very framework of drives they make possible.

Returning to the car with its trunk open is essential to highlight the abductive superiority of Nietzsche's model of mind over the Socratic. The trunk example I used to explain abduction is relevant here because there are other causes for trunks to open which do not depend on intentionality. Trunks may open for no apparent reason, at least if intentionally construed, if the Body Control Module (BCM) that manages the car's electronic locking sensor is faulty. An individual with expert knowledge of such modules (a quality control automobile engineer, for example) and a familiarity with the car's make, model, and approximate year in question (which the engineer is likely to know) might reason via abduction very differently than the average person if he observed the same car with its trunk open in a parking lot. Based on his expertise and the information about the car's year and model, he may conclude that the BCM is the cause of the trunk opening and needs to be replaced. Most people would assume that the expert's hypothesis is more warranted than the "average Joe's."

Nietzsche's abductive postulation is analogous to this second expertly informed explanation of the car's trunk failing to remain closed. According to Nietzsche's psychological model, so-called intentionality is not a causal factor in generating action. At best, an agent's supposed 'intentionality' at any given time is informed and colored by the active drive. Nietzsche, of course, had a distinct advantage over Socrates, much like the car engineer in the above example, because he was well-versed in his day's biological, anthropological, and physiological disciplines. These studies revealed a set of physical systems and set-ups that evolved to platform conscious and unconscious behavior. As noted above, Nietzsche's postulations are viable experimental models that may bear fruit with advances in neuroscience. In contrast to Stern, Nietzsche's drive theory is preferable to Socrates's psychological model because its abductive postulation is testable and coheres with verifiable contemporary neuroscientific hypotheses. It is not explanatorily idle, as Stern mistakenly contends.

In conclusion, both Nietzsche and Socrates understand the problem of directly justifying the structures that scaffold human psychology. For Nietzsche, drives are the ultimate building blocks of all consciousness: they serve as engines for action, motivate our interests, and frame our experiences. If Nietzsche is correct, then drives are beyond direct confirmation because perception is never value-

neutral. As Nietzsche rightly stated: "Facts are precisely what there is not, only interpretations." (*Will to Power* 481).

For Nietzsche, drives have three aspects: they are transformed animal instincts, they run in fixed channels, and on the mental side, are expressed as intense distinctive urges signifying to their bearers what needs to be done to fulfill the particular urge. The actual source behind these urges cannot be directly known: we cannot always fathom the real reason for the drive's expression. Thus, in some sense, we might say that drives act as Kantian transcendental conditions for the very possibility of action and knowledge. However, rather than speaking about drives in Kantian language, Nietzsche postulates them as naturalistically inflected hypotheses. Our anthropological, physiological, psychological, and historical investigations reflect them. These studies, much like Socrates's dialectic, give us greater, albeit indirect, insight into the nature and purpose of drives.

For Socrates, Forms take the place of animal instincts that become transformed into drives. They, too, are abductive postulates necessary, according to Socrates, for platforming discussions of knowledge and moral action. Socrates's psychology is divided into *elenchus* and access to the Forms. The *daimon* is an inverted version of the propulsive force of Nietzsche's drives as instead of motivating a subject to perform some action, it arrests the flow of an agent's movement and thinking, forcing him to reconsider and reinvestigate his reasons for some conclusion. Again, like the urge aspect of drives, the *daimon* is the bridge linking the Forms to *elenchus*. Like drives, Forms cannot be directly viewed; doing so would be like looking at the sun. Yet the Socratic dialectic can give a better sense of what each Form of virtue entails, much as drives also have distinct channels that allow them to be expressed. A proper definition acts as a reflective mirror on the Forms, illuminating aspects of virtues that we would not otherwise know.

Stern's overall diagnosis is incorrect both in form and substance. Neither Socrates nor Nietzsche believe they provide representative components of human psychology predicated on direct evidence. Their arguments are abductive in form. Furthermore, Stern does not fully appreciate the complexity of Socrates's model and, therefore, fails to understand the relationship between its fundamental components: Forms, Daimon, and *elenchus*. I have fleshed out the substantive relationships of each component here.

The significance of comparing the two models is now clear: since both rely on abductive hypotheses as the cornerstone of their respective models of human psychology, the principal token of Socrates's philosophical psychology, namely the process of deliberation, is necessarily unknown to itself: deliberation cannot see what scaffolds it. Such a realization gives us reason to prefer Nietzsche's model



since it is more honest about its blindspots and is in keeping with a contemporary scientific view of the mind, which recognizes the brain in particular and the body more generally as the biological underpinning of 'reason.' Most significant, however, is Nietzsche's inversion of Socrates's philosophical psychology. Far from being antithetical, it is more correct to say that Nietzsche's drive theory marks a significant improvement over Socratic philosophical psychology. Indeed, drive theory is a correction of the theory of the Forms, which explains why Socrates is always "close" to Nietzsche.

### Bibliography

ALFANO, Mark. *Nietzsche's Moral Psychology*, Cambridge University Press, 2019.

ARAÚJO, H. F. de; MONTENEGRO, M. A. de P. "Perception and Memory in Plato." *Kalagatos*, [S. l.], v. 21, n. 2, p. eK24045, 2024. Disponível em:  
<https://revistas.uece.br/index.php/kalagatos/article/view/13376>. Acesso em: 13 jul. 2024

CLARK, Maudmarie and DUDRICK, David. *The Soul of Nietzsche's Beyond Good and Evil*, Cambridge, Cambridge University Press, 2012.

DEVEREAUX, Daniel, T. "The Unity of the Virtues in Plato's Protagoras and Laches", *The Philosophical Review*, Vol. 101, No. 4 (Oct., 1992), pp. 765-789.

DOUVEN, Igor, "Abduction", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL =  
<<https://plato.stanford.edu/archives/sum2021/entries/abduction>>

HUBY, Pamela. "Phaedo 99 D-102 A" *Phronesis*, Vol. 4, No. 1 (1959), pp. 12-14

JETLI, P. "Deduction–Abduction–Induction Chains in Plato's Phaedo and Parmenides", In: Magnani, L. (eds) *Handbook of Abductive Cognition*. Springer, Cham.

KATSAFANAS, Paul. "Nietzsche's Philosophical Psychology", in K. Gemes and J. Richardson (eds.), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press), 2013, 726–55.

KATSAFANAS, Paul. *The Nietzschean Self: Moral Psychology, Agency and the Unconscious*, New York: Oxford University Press, 2016,

KAUFMANN, Walter. "Nietzsche's Admiration of Socrates" *Journal of the History of Ideas*, 1948, 472-491.

- KRAUT, R. 1981. "Plato's Apology and Crito: Two Recent Studies" *Ethics*, 91: 651-664
- McPherran, M. 1996. *The Religion of Socrates*. University Park: The Pennsylvania State University Press
- LEITER, Brian. "The Paradox of Fatalism and Self-Creation in Nietzsche," in *Willing and Nothingness: Schopenhauer as Nietzsche's Educator*, ed. Christopher Janaway Oxford: Clarendon, 1998.
- LIGHTBODY, Brian. *The Problem of Naturalism*. Lanham MD: Lexington Press, 2013.
- LIGHTBODY, Brian. *Dispensing the Clouds of Temptation*. Eugene Oregon, Pickwick Press, 2015.
- LIGHTBODY Brian. *Nietzsche's Will to Power Naturalized: Translating the Human into Nature and Nature into the Human*. Lanham Maryland: Rowman and Littlefield, Lexington Books, 2017.
- LIGHTBODY, Brian. "Twilight of the Genealogy or A Genealogy of Twilight? Saving Nietzsche's Internalization Hypothesis from Naïve Determinism." *Philosophical Readings*, Volume 3, Sept 2021 183-194.
- LIGHTBODY, Brian. "The Passive Body and States of Nature: An Examination of the Methodological Role State of Nature Theory Plays in Williams and Nietzsche." Special Issue: Philosophical Genealogy from Nietzsche to Williams *Genealogy* 5 (2), 38, 2021, 1-15.
- LIGHTBODY, Brian. *A Genealogical Analysis of Nietzschean Drive Theory*. London: U.K. Palgrave Macmillan, 2023.
- LIGHTBODY Brian. "On Becoming Fearful Quickly: A Reinterpretation of Aristotle's Somatic Model of Socratean Akrasia." *The Journal of Ancient Philosophy*, Vol. 17 Issue 2, 2023, 134-161.
- LISBOA, Helena M. O "inconsciente e a psicologia de Nietzsche. (The unconscious and Nietzsche's Psychology) *Kalagatos*, [S. l.], v. 11, n. 22, p. 383-399, 2021. Disponível em: <https://revistas.uece.br/index.php/kalagatos/article/view/6134>. Acesso em: 13 jul. 2024.
- MARCUSE, Herbert. *Eros and Civilization*, 2nd edition. Boston: Beacon P, 1966.
- McPHERRAN, M. *The Religion of Socrates*. University Park: The Pennsylvania State University Press, 1996.
- MELE, Alfred. "Akrasia, "Self-Control, and Second-Order Desires", *Noûs* Vol. 26, No. 3 Sep.1992.
- MOTA, T. "Intuição do caos e inversão do ônus da prova: Perspectivismo e filosofia experimental em Nietzsche. [INTUITION OF CHAOS AND REVERSAL OF THE CHARGE OF PROOF:PERSPECTIVISM AND EXPERIMENTAL PHILOSOPHY IN NIETZSCHE] *Kalagatos*, [S. l.], v. 17, n. 1, p. 49-61, 2021. Disponível em: <https://revistas.uece.br/index.php/kalagatos/article/view/6454>. Acesso em: 13 jul. 2024.
- NIETZSCHE, Friedrich. *Beyond Good and Evil in Basic Writings of Nietzsche*. Trans. Walter Kaufmann. Intro. Peter Gay. New York: Random House, 2000

NIETZSCHE, Friedrich. *Human all too Human*. Trans. R.J. Holingdale. Intro by Richard Schacht. Cambridge University Press: 1996.

NIETZSCHE, Friedrich. *The Gay Science*. Trans. Walter Kaufmann. New York: Vintage Books, 1974.

NIETZSCHE, Friedrich. *On the Genealogy of Morals in Basic Writings of Nietzsche*. Trans. Walter Kaufmann. Intro Peter Gay New York: Random House, 2000.

NIETZSCHE, Friedrich. *Twilight of the Idols/The Anti-Christ*. Trans. R. J. Hollingdale. Harmondsworth, Middlesex: Penguin, 1972.

NIETZSCHE, Friedrich. *Ecce Homo*. Trans Walter Kaufmann, *Basic Writings of Nietzsche*, New York: 2000.

NIETZSCHE, Friedrich, *The Will to Power*. Trans Walter Kaufmann and R.J. Holingdale. Vintage Books 1970.

NUSSBAUM, Martha. "Introducing a New God: Socrates and His Daimonion Commentary on Edmunds." *Proceedings of the Boston Area Colloquium of Ancient Philosophy* 1: 231-240, 1985.

PANKSEPP, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York, NY: Oxford University Press.

PANKSEPP, J. (2011). Cross-species affective neuroscience decoding of the primal affective experiences of humans and related animals. *PLoS ONE* 6:e21236. doi: 10.1371/journal.pone.0021236

PANKSEPP, J., NORMANSELL, L., COX, J., and Siviy, S. (1994). "Effects of neonatal decortication on the social play of juvenile rats." *Physiol. Behav.* 56, 429–443. doi: 10.1016/0031-9384(94)90285-2

PANKSEPP, J. B., and HUBER, R. (2004). Ethological analyses of crayfish behavior: a new invertebrate system for measuring the rewarding properties of psychostimulants. *Behav. Brain Res.* 153, 171–180. doi: 10.1016/j.bbr.2003.11.014

PANKSEPP, J., and YOVELL, Y. (2014). Preclinical modeling of primal emotional affects (Seeking, Panic and Play): gateways to the development of new treatments for depression. *Psychopathology* 47, 383–393.

PLATO. *Apology*. *Plato the Complete Works* ed. John M. Cooper Trans. G.M.A Grube, Indianapolis: Hackett Publishing 1997

PLATO. *Euthyphro*. *Plato the Complete Works* ed. John M. Cooper Trans. G.M.A Grube, Indianapolis: Hackett Publishing 1997.

PLATO. *Gorgias*. *Plato the Complete Works* ed. John M. Cooper Trans. Donald Zeyl. . Indianapolis: Hackett Publishing 1997.

PLATO, *Phaedo* in *Plato the Complete Works*, ed. John M. Cooper, trans. Stanley Lombardo and Karen Bell Indianapolis: Hackett Publishing, 1997.

PLATO. *Republic* in *Plato the Complete Works* ed. John M. Cooper Trans. G.M.A. Grube. Indianapolis: Hackett Publishing 1997.

RICCARDI, Mattia. "Virtuous Homunculi: Nietzsche on the Order of Drives", *Inquiry an International Journal of Philosophy*, 61 (1) 21-41 2018

RICCARDI, Mattia. *Nietzsche's Philosophical Psychology* Oxford: Oxford University Press, 2021.

RICHARDSON, John. *Nietzsche's New Darwinism*. Oxford University Press, 2004.

ROBINSON, Richard *Plato's Earlier Dialectic* (Ithaca, New York: Cornell University Press, 1941). 56- 57.

ROBINSON, Tim. "Plato's Separation of Reason and Desire," *Phronesis*, 16 (1971) 38-48.

SELLARS, Wilfrid. *Science, Perception and Reality* London: Routledge and Kegan Paul, 1963.

R.E. Stalley, R.E. "Plato's Argument for the Division of the Reasoning and Appetitive Elements within the Soul", *Phronesis* 5 20.2 (1975) 110-128

STERN, Tom. "Against Nietzsche's 'Theory' of Drives" *Journal of the American Philosophical Association* Vol 1. Issue 1 Spring (2015) 121-140.

STOVALL, Preston. *The Single-Minded Animal: Shared Intentionality, Normativity, and the Foundations of Discursive Cognition*, New York: Routledge, 2021

VAN RIEL, Gerd. "Socrates' Daemon: Internalisation of the Divine and Knowledge of the Self", *Apeiron: A Journal for Ancient Philosophy and Science* , June 2005, Vol. 38, No. 2 31-42, 38.

WELSHON, Rex. *Nietzsche's Dynamic Metapsychology: This Uncanny Animal* New York: Palgrave 2014.



LIGHTBODY, Brian; COSTA, Pâmela. Drives as Inverted Forms: Nietzsche's Correction of Socrates's Philosophical Psychology. *Kalagatos*, Fortaleza, vol. 21, n. 2, 2024, eK24047, p. 01-29.

Received: 06/2024

Approved: 07/2024