Univerzita Karlova

Filozofická fakulta

Ústav filosofie a religionistiky

Diplomová práce

Bc. Filip Váňa

Beyond Psychology:
Nietzsche and the Metaphysics of Agency

Za hranice psychologie:
Nietzsche a metafyzika jednání

Praha 2022

Vedoucí práce: Mgr. Tereza Matějčková, Ph.D.

Acknowledgment:

This work would not have been what it is without many people. First of all, I would like to thank Dr Tereza Matějčková for her encouraging supervision of this thesis and for all the trust she has shown me throughout my studies. I also thank Dr Alexander Prescott-Couch for the many fruitful meetings we had during my stay at Oxford. Mikołaj Sławkowski-Rode and Ralph Weir have my thanks for making that stay and research fellowship possible, but also for all the inspiring conversations we had shared during our meetings in Warsaw, Zagreb, and Oxford. Finally, when it comes to philosophy itself, there is nobody among the living to whom I am more indebted than Jensen Suther and Martin Hägglund. Despite being on the other side of the Atlantic, our reading group and our ongoing discussions and email exchanges have shaped my thinking more than anything over the past year. Thank you.

Declaration:

I hereby declare that I have written this Master's Thesis independently, using only the mentioned and duly cited sources and literature, and that the work has not been used in another university study programme or to obtain the same or another academic title.

In Prague on the 4th of August 2022 Bc. Filip Váňa

Key Words:

Metaphysics of Agency, Logic of Agency, Nietzsche, Psychology, Life, Drive, Organism, Consciousness, Reason, Freedom

Klíčová slova:

Metafyzika jednání, Logika jednání, Nietzsche, Psychologie, Život, Pud, Organismus, Vědomí, Rozum, Svoboda

Abstract:

In this work I attempt a reconstruction of what I take to be the very foundations of Nietzsche's philosophical project – his account of what it means to be an agent (i.e. his "psychology"). Yet, my main purpose is far from mere exegesis. On the contrary, by actively thinking through what Nietzsche gives us, I aim to develop a coherent account of the fundamental "logic" of agency as such. That is, of what it could mean to be an agent, understood as a functionally distinct kind of being. In doing so, I proceed in a number of steps. First, I situate Nietzsche's project in the historical context and say a few words about what he has to offer us. Next, I consider Nietzsche's approach to how we should think about the soul (and psychology itself) and attempt to clarify some other key psychological concepts. With these in hand, I develop a formal concept of an organism and connect it to Nietzsche's notorious notion of the "will to power" – which concludes the first part of the work. In Part Two, the "psychological" account worked out in Part One meets ethics. Here I begin by examining Nietzsche's (internalist) theory of motivation, itself derived from the concept of life, and scrutinize his attempted deconstruction of the "selfish/selfless" dichotomy – the foundation of Nietzsche's "moral psychology" and the basis for his rejection of traditional morality. I argue that Nietzsche's initial strategy fails, but I also show that he himself soon opts for a different one – namely, one that challenges our rationality, our capacity to set our own ends (and know "why we're doing what we're doing"), and thus threatens to reduce agency itself to a subjective fiction. However, in the final chapter, I demonstrate that this challenge proves incoherent and must be rejected. I conclude the work by discussing why a merely psychological account of agency cannot suffice; and why it proves necessary to venture "beyond psychology" – into critical metaphysics.

Abstrakt:

V této práci se pokouším o rekonstrukci toho, co považuji za základ Nietzscheho filosofického projektu – jeho koncepci jednání (jeho „psychologii“). Hlavní účel této práce však není ani zdaleka exegetický. Naopak, na základě aktivního promýšlení Nietzscheho pestrého díla se snažím vypracovat ucelené pojetí fundamentální „logiky“ jednání vůbec. Tedy toho, co to znamená být aktérem, jakožto specifickým druhem aktivní, jednající bytosti. V práci postupuji v několika krocích. Nejprve Nietzscheho projekt zasadím do historického kontextu a řeknu pár slov o tom, co nám může nabídnout. Dále zvážím Nietzscheho přístup k tomu, jak bychom měli uvažovat o duši (a psychologii samotné) a pokusím se objasnit jeho klíčové psychologické pojmy. První část práce uzavřu vypracováním formálního konceptu organismu a jeho propojením s Nietzscheho pojmem „vůle k moci“. Ve druhé části se pak „psychologie“ promyšlená v části první střetne s etikou. Nejprve se zaměřím na Nietzscheho (internalistickou) teorii motivace, vycházející právě ze samotného pojetí života, a pečlivě promyslím jeho pokus o dekonstrukci dichotomie sobeckosti a nezištnosti, jež má tvořit základ Nietzscheho „morální psychologie“ a jeho zavržení tradiční morálky. Zde dojdu k závěru, že Nietzscheho úvodní kritika morálky selhává, zároveň však předvedu, že Nietzsche sám následně volí strategii jinou – a to takovou, jež staví na kompletním zpochybnění naší racionality a schopnosti stanovovat si účely vlastního jednání (a tedy vědět, „proč děláme, co děláme“). Demonstruji, že pokud by tato strategie měla být úspěšná, musela by popřít i samotnou realitu jednání vůbec. V závěrečné kapitole ji tak promýšlím – a docházím k tomu, že se jedná o strategii ryze nekoherentní. Práci následně uzavírám diskuzí o tom, proč pouze psychologická koncepce jednání nemůže stačit; a proč se je třeba vydat „za hranice psychologie“ – do domény kritické metafyziky.

Beyond Psychology

Nietzsche and the Metaphysics of Agency

Contents

List of Abbreviations

Nietzsche's works are cited by using the conventional abbreviations listed bellow. Roman numerals refer to major parts or chapters; Arabic numerals to the section numbers, not pages. In writing this work, I have consulted the German critical edition (KSA), as well as multiple English translations. For complete information, see the Bibliography at the end.

A	*The Antichrist*
BGE	*Beyond Good and Evil*
BT	*The Birth of Tragedy*
D	*Daybreak*
GM	*On the Genealogy of Morality*
GS	*The Gay Science*
HH	*Human, All Too Human*
KSA	*Sämtliche Werke: Kritische Studienausgabe in 15 Bänden*
TI	*Twilight of the Idols*
TL	*On Truth and Lying in Non-Moral Sense*
WP	*The Will to Power*
Z	*Thus Spoke Zarathustra*

Introduction

What a piece of work is a man! How noble in

reason, how infinite in faculty! In form and moving how

express and admirable! In action how like an angel, in

apprehension how like a god![1]

~~~

Friedrich Nietzsche's philosophical project as a whole, despite its varied and fragmentary character, remains thoroughly committed to what Nietzsche himself called "translating man back into nature".[2] Not at all in the naive sense of giving up on culture, the fruits of civilization, and returning back to the nomadic life or some fantastical "state of nature", but rather in a sense of profound *naturalism*: A general philosophical framework that rejects all transcendent, otherworldly, esoteric, mystical, or otherwise mysterious, that is, *supernatural*, explanations – for they are, of course, no explanations at all – and instead endeavours to understand the diverse phenomena we encounter in our lives, be they ordinary or highly extraordinary, by interpreting them, so to speak, from the inside.[3] All the while insisting on *no fundamental gaps*.

Today, the supernatural, theocentric picture of the universe, which has animated mankind's self-conception for millennia, is mostly gone. For the most part, we no longer think of ourselves as creatures of God, endowed with divine powers, no matter how fallen. Instead, the task has

1 Shakespeare W., *Hamlet*, II.2.303–307.

2 BGE 230.

3 BGE 36.
~~~

become to understand ourselves as merely another kind of animal, as a product of evolution, of natural history. As bearing not the image of God, but rather "the indelible stamp" of our "lowly origin".[4] But already in Nietzsche's own time the old world was beginning to wither away, beliefs supporting it "becoming unbelievable".[5] And even according to Nietzsche himself, the new, naturalistic world-view – one that remains true to *this world* – was finally (at least in some parts of the world) becoming reality. Or perhaps we should say, a real possibility.[6] For Nietzsche thought this epochal task very far from finished indeed. Much of the anxiety, frustration, and indignation, but also hope, auspiciousness, and "cheerfulness",[7] so characteristic of Nietzsche's voice, are due to Nietzsche's sense of his own untimeliness, of finding himself "stretched in the contradiction between today and tomorrow",[8] a historic interregnum where the old is dying but the new cannot yet be born.

> "*New struggles* (...) God is dead: but the way humans are, there may exist caves in which they show his shadow for millennia yet. – And we – we have still to defeat his shadow!"[9]

Of course, one rarely thinks of Nietzsche as a philosopher of nature. Even superficial skimming of his writing reveals his main preoccupations and interests lie elsewhere – in the *human* sphere. Yet his sober naturalism remains a general philosophical (methodological, ontological,

---

4 Darwin, C., *The Descent of Man, and Selection in Relation to Sex*, Princeton University Press, 1981, p. 405.

5 GS 343.

6 Nietzsche never tired of pointing out, with contemptuous amazement, that despite the momentous shift the Death of God should signify all around God's shadows continue to hold sway over people's lives as if nothing major happened; and that consequently, most modern unbelievers – seeing as they unthinkingly accept, even cling to, the old values, ignorant of both their origins and contingency – effectively remain merely diluted, secularized Christians. As a result, because Nietzsche took the notion of God seriously, he treated most self-avowed "atheists" with contempt.

7 See GS 343.

8 *Ibid*.

9 GS 108.

cosmological) framework and attitude that constrains and shapes the rest; a prism through which we must look, if we want to understand Nietzsche *at all*. This has not been lost on the modern scholarship, and it is one of the main reasons why Nietzsche remains a relevant thinker today.[10] For it is often precisely such naturalistic commitments that lead Nietzsche to ask what this new age with its newfound "joyful wisdom" would come to usher amongst ourselves; what it would do *to ourselves*.

> "When shall all these shadows of God cease obscuring us? When shall we have nature utterly undeified! When shall we be permitted to *naturalize* ourselves in terms of the pure, newly discovered, newly redeemed nature!"[11]

Just who exactly are we? How did we get here? What does it mean to live in a Godless world "beyond good and evil"? What shall we make of the traditional human privileges if we are but another step of an immanent natural process? How will this tectonic shift alter our perceived place in the world? Our self-conception? Our societies? *These* are Nietzsche's questions – and the reason why one sooner thinks of Nietzsche as an ethical thinker, a philosopher of culture, religion, art, and perhaps most important of all, a kind of "psychologist". One who, in wrestling with these questions, comes to conclude that large amounts of the "received wisdom" are, quite simply, based on an error.

And at the very centre of those errors, one finds the concept of *agency* itself. Perhaps the deepest, most fundamental of all Nietzsche's targets, even if not the most explicit one. The one

---

10 For instance, it is worth noting that Nietzsche was among the first philosophers who took evolution and heredity (that is, natural history, roughly as we now understand it) *truly seriously* (despite his frequent criticisms of "Darwinism").

11 GS 109.

notion that envelops and concentrates just about every topic Nietzsche's thought touches. Motivation, responsibility, freedom, rationality, instincts, human "nature", temperament, creativity, soul and body, activity and passivity, strength and weakness, power and suffering, value, purpose, or life. To name a few.

What better concept, then, to take as *our* target? And who better than Nietzsche – an unflinching naturalist who attempts its radical revision, a reevaluation of our self-conception, of our understanding ourselves as rather exceptional kinds of agents (and even perhaps, agents *as such*) – to think it through with? Let us then delve into Nietzsche's thought and attempt to bring this concept out onto the surface, to make it explicit, and in the process understand what Nietzsche thinks – as well as what *we should think*.

~~~

To get a general picture of what is to follow, as well as of some of the more specific reasons for pursuing this topic this way, let me begin by considering the question of what it is that Nietzsche really takes issue with. What about the traditional understanding of agency does he believe untenable, and why?

I say "traditional understanding of agency". But is there such a thing? Perhaps surprisingly (then again, perhaps not), at least with regard to *human* agency – when it comes to the question of what makes us the kind of agents we are, or really, the kinds of beings we are – one notices rather little historical variety. In fact, we really only see one general picture:[12] A shared

12 Of course, when we look closely at, say, the somewhat more specific issue of free will (i.e. how to reconcile free agency with some perceived obstacle, be it deterministic laws of nature or God's foreknowledge) different theories start appearing. In fact, to an average modern reader, steeped in the voluntarist language of unencumbered "choice", the classical account of freedom *qua* rational agency (where free is not one who can choose, but one who chooses *well*), probably sounds barely comprehensible. Even so, the essential outline, the
~~~

inheritance of pagan antiquity and medieval Christianity, which has always grounded man's essence (and uniqueness) in his rationality and freedom. On this picture, in contrast to animals that rely on their *reflexes*, as humans we can also *reflect* on things, gain distance on our immediate motives, decide our own ends, and achieve fuller "control" over what we do. In broad strokes, this remains the "common-sense" picture today. But of course, we don't really think other animals are merely inert "things" either. They are living beings; *agents* in their own right. By which we usually mean: a being or an "entity" capable of some form of self-determination. Which is to say, a beings who can, at least in some capacity, *act*, rather than merely passively react. A being "things can be up to" (rather than up to its properties or some events occurring inside it), capable of "settling things" on its own, capable of *having and leading a life.*

To Nietzsche the naturalist, however, this poses a serious problem. For just *how* could, in a world where there can be no "endowing", self-determining and thinking agents ever emerge? How could unreason and inert matter give rise to their opposites? Nietzsche often criticizes other philosophers for their lack of "historical sense": for prioritizing "being" over "becoming" and failing to appreciate that the real origin of a thing is often completely different from the status we later ascribe to it. And he also remarks that our basic philosophical question is the same it's always been: How can a thing arise from something other than itself?[13] This is what he has to say:

---

intimate connection between reason and freedom, remains preserved throughout.

13 See HH I.1: "Philosophical problems today take on the same form in almost every respect as they did two thousand years ago: how can something arise from its contrary – for example, reason from unreason, sensitivity from lifelessness, logic from illogic, disinterested contemplation from covetous desire, selflessness from egoism, truth from error?"

> "Until now, metaphysical philosophy has overcome this difficulty by denying the origin of the one from the other, and by assuming for the more highly valued things some miraculous origin, directly from out of the heart and essence of the 'thing in itself' (...) Historical philosophy, on the other hand, which can no longer be thought as separate from natural science, the most recent of all philosophical methods, has determined (...) that they are no opposites, except as exaggerated by the popular or metaphysical conception, and that this contrast is based on an error of reason."[14]

Thus, at the very lest, it would appear that the things we've always taken for granted and valued so highly – be it freedom, reason, agency, or life – cannot very well be what we took them for. Accordingly, if we are to keep using these terms at all, their content – as well as their value – will have to be radically rethought.

Overall, these are the main reasons why Nietzsche's naturalism, along with his critique and account of agency (the "psychology"), ought to be understood as absolutely foundational to his thought. The other labels mentioned above, such as that of a critic of morality, might capture what Nietzsche was after, but only these can help us understand how he got there. They are the grounds, even the "method", from which the rest – the main project revolving around the "transvaluation of values" and the critique of morality entailed in it – follows.[15]

---

14 HH I.1.

15 But is Nietzsche truly an "objective" thinker? One concerned with truth, who lets his subject matter lead him where it will? Or is he more concerned with achieving his preferred practical ends (a noble, aristocratic culture unafraid of greatness, capable of producing great individuals, great art, and so on)? This internal strife is something of a Nietzsche's signature. One one hand we get Nietzsche the thinker: someone who believes to have discovered important truths, even glimpsed the future. But on the other, there is Nietzsche the man: an individual of flesh and blood, terrified by what he sees coming – and who thus tries to combat it with all his (sometimes deceptive) might. Brian Leiter has usefully distinguished between the "Humean Nietzsche" and the "Therapeutic Nietzsche". Yet Leiter overestimates the degree to which we can neatly distinguish between the two in practice. Or rather, underestimates how the revolutionary or "therapeutic" project sometimes infects and distorts the theoretical one. That being said, we can, of course, evaluate Nietzsche's arguments by judging them on their own merits. And that is all that matters to me here.

This makes Nietzsche a member of the long line of thinkers we today like to call "moral psychologists". That is, thinkers whose conception of human nature (in the broadest sense) directly informs their practical philosophy. As Nietzsche sees it, however, the other members of this club have generally got things the other way around. They let morality, the "Circe of philosophers",[16] contaminate their accounts of humans and humanity, as well as nature itself, only to end up with a wishful, prejudiced, moralized, and distorted conception of all three: nature, agency, *and* morality. The result of this, at least in Nietzsche's eyes, is a picture so full of holes and metaphysical fictions (themselves but slightly more sophisticated versions of the primitive superstitions and constructions of religion)[17] that it is perhaps not so much in need of revision as of retirement – and, of course, replacement. And naturally, Nietzsche would have us think that he is here to supply us with just that. His cold, naturalistic psychology – "the path to fundamental problems"[18] – exposing just what these fictions are in truth. The question we must ask is therefore simple: Just how credible of an account is it?

The answer I give and argue for bellow is ... equivocal. Almost invariably, Nietzsche unearths and point to decisive problems, and many of his insights prove valuable in their own right. Yet, in the end, once all the part are in place, I contend that the "Nietzschean picture" manifestly fails – cannot hold together because it is undermined by Nietzsche's sweeping scepticism. To show all this, I attempt a reconstruction of these essential elements, which, when put together, yield what I take to be the most significant aspect of Nietzsche's philosophy as such. That being said, my goals are far from exegetical. On the contrary, by actively thinking

16 D 3.

17 See for example A 14.

18 BGE 23.

through what Nietzsche gives us (and in dialogue with the recent scholarship), I aim to develop a coherent account of the *fundamental structure of agency* as such – of what it means to be an agent, understood as a functionally distinct kind of being: a being that undergoes subjective experience and is capable of some form of self-determination. As such, this work can be seen as an exercise in what I call the "metaphysics of agency", the point of which is to establish the necessary "logic", or what must be true, of any possible "agential" experience in general.

I proceed in a number of steps. First I take up some of the things mentioned in the Introduction and draw a few historical connections. Most importantly, I point out two decisive features of Nietzsche's thought: the fact that he inherits modern philosophy's "subjectivism", but repudiates its "rationalism". This leads me to ask what, on Nietzsche's account, subjectivity (or subjects) and rationality (reason) actually amount to. In a way, this can be understood as the main question running throughout this entire work. In the second chapter, I begin answering it by clarifying the role of some of the main "psychological" concepts Nietzsche uses, such as the conscious and the unconscious. I also consider Nietzsche's approach to how we should think of the soul (and psychology itself). And from there, it is only a short way to the centrepiece of Nietzsche's picture, his notion of *the drive*. After its introduction, in Chapter III, the concept is subjected to careful scrutiny, and although it is found seriously lacking (or underdeveloped) in certain ways, it nevertheless proves a valuable component for our general account of agency as such. Subsequently, by putting these components together, the fourth chapter presents a formal concept of the *organism*: an irreducibly normative, internally purposive being – an agent – that has (and strives toward) a "good". Finally, I connect this "conative" dynamic of living to Nietzsche's notorious notion of the "will to power", and this brings Part One to a close.

In Part Two, the conception of agency and life (the "psychology") worked out in the first part finally meets *ethics*, or moral inquiry. In the fifth chapter, I examine Nietzsche's (internalist) theory of motivation, itself derived from the very notion of life as will to power, and critique his attempted deconstruction of the "selfish/selfless" dichotomy – the foundation of Nietzsche's "moral psychology" and the basis for his rejection of traditional morality. In the sixth chapter, I argue that Nietzsche's initial strategy fails, but I also show that he has another one; one grounded in a rather strong scepticism about self-conscious reason as such: our ability to set our own ends and to truly know "why we're doing what we're doing". I demonstrate that this strategy, if successful, destroys everything significant we've been through and reduces agency to a mere fiction. However, in the final, seventh chapter, I come to conclude that even this strategy must be rejected, as it proves to stand on highly untenable, often incoherent assumptions. Accordingly, it will become clear why a merely psychological account of ourselves cannot suffice; and why we must instead venture "beyond psychology" – into critical metaphysics.

Part One

I. Subjectivism and Rationalism

The one universal charge Nietzsche levels against all previous ("metaphysical") philosophy is, to put in his idiom, that of *bad psychology* – a psychology "stuck on moral prejudices and fears" that "has not dared venture out into the depths".[19] So Nietzsche tells us and urges the psychologists of the future to wage war against such prejudices, against their heart's desires and all forms of moralizing, so that psychology – the "path to the fundamental problems"[20] – may be rightfully recognized as "the queen of the sciences".[21]

Among other things, this brief "profile", which will be further developed in what is to come, brings us to the uneasy question of how to situate Nietzsche's project historically, as well as how to properly understand Nietzsche's own conception of it (what he actually took himself to be doing). Obviously, a comprehensive account of Nietzsche's relationship to the history of thought cannot be attempted here. However, the elements of Nietzsche's project most relevant to this work are obviously best understood in context. Let us therefore attempt to draw them out as such.

One important ingredient was already discussed in the Introduction. Namely, Nietzsche's "naturalism": a philosophical "attitude" (if not a metaphysics in its own right) closely related to the birth of modern natural science and decisive for modern thought in general. But there are other important strains of "modern" thought that will prove crucial. Now Nietzsche

19 BGE 23

20 *Ibid*.

21 *Ibid.*

relationship with modernity, and especially the so called Enlightenment, doesn't admit of simple treatment either. But the question I want to ask now is precisely: *Why?* Which key parts of modern philosophy does Nietzsche reject and accept? If we can answer that question – and I believe we can – then, I propose, we will have gained a crucial component for understanding the essence of Nietzsche's basic project.

On one hand, we might think that Nietzsche the naturalist should be very much on Enlightenment's side. After all, doesn't the Enlightenment stand precisely for a radical break with Tradition? Doesn't it embody the revolutionary *Zeitgeist* of a thorough "disenchantment of the world"? And shouldn't Nietzsche welcome these things? No doubt. As I observed in the Introduction, the problem here seems to be mainly that, in Nietzsche's eyes, that disenchanting remains woefully incomplete. God might have been killed, but his shadows dance on. – And the largest one of them all (and according to Nietzsche, the most harmful): our traditional picture of the *agent*; that is, the *rational* and *moral* agent. Indeed, despite the revolutionary spirit, modern philosophy's treatment of (human) agency – from Descartes to Hegel (apart from a few exceptions)[22] – must have seemed to Nietzsche merely a faithful successor to the millennia-long Platonic-Christian legacy. In fact, Nietzsche often treats this modern instalment as the most perverse version of it yet. Why?

With Descartes, the question of human nature, and of the human Ψυχή, gets posed anew. And when the classical tripartite soul emerges from under Descartes' knife, all that's left is nothing but its last, rational part. The pure "I" as the disembodied, self-sufficient, self-transparent, and self-conscious *thinking subject*, whose self-certainty becomes the ultimate

22 Most notably Hume, but also Spinoza, Pascal, or La Rouchefoucauld.

measure and guarantor of truth. Now much of this – such as disembodiment or self-transparency – Nietzsche has no patience with. And yet, it is here where we should search for the germ of Nietzsche's ambivalence toward modern philosophy and the heart of his own project. The two key words are: *thinking* – *subject*. As I suggest and briefly argue in the remainder of this chapter – although the real argument is really only this work as a whole – Nietzsche's philosophy must be understood as animated by these two concepts.

When Descartes in his quest for certainty retreats into himself and leaves the external world behind, something strange happens. For when he starts putting the world back together, reconstructing it piece by piece from the indubitable ideas he finds in himself, what emerges is not quite the original. In fact, it is its *copy*: the world as a subjective "*re*-presentation" of what is, supposedly, objectively "present". Thus the world becomes My Idea. Or to put it in Heidegger's idiom: with the "subject" Descartes ushers in "the age of the world picture" – and sets the stage for modern metaphysics that culminates only with ... Nietzsche. Nietzsche, the one in whom, according to Heidegger's diagnosis, this whole pathological tradition of modern metaphysics, and in fact the whole age, finds its "truth".

> "No matter how sharply Nietzsche pits himself against Descartes, whose philosophy grounds modern metaphysics, he turns against Descartes only because the latter *still* does *not* posit man as *subiectum* in a way that is complete and decisive enough. (...) Modern metaphysics first comes to the full and final determination of its essence in the doctrine of the Overman, the doctrine of man's absolute pre-eminence among beings. In that doctrine, Descartes celebrates his supreme triumph."[23]

23 Heidegger M., *Nietzsche*, Volume IV, HarperCollins, 1991, p. 28.

On Heidegger's reading, subjective re-presentation of the world of objects is not good enough. After all, Descartes, like most of his heirs, still believed in the highest being, the Christian God, creator of both man and nature, and the ultimate, objective, and absolute "value": the True, the Good, the Beautiful – as the ontological fullness of creation, and the transcendental end of all desire. No, it is only with the *Übermensch*, the apotheosis of the Nietzschean ideal of the *self-creation of values,* that man himself becomes the real measure of things and the centre and aim of all beings. And it is thus here that modern philosophy finally culminates – as instrumentalising, objectifying, anthropocentric. As subjectivism that makes the world and everything in it (other humans not excluded) *our object* – the "standing reserve" awaiting to be used and exploited by the human will. Or so Heidegger tells us.

But now is obviously not the time for careful evaluation of Heidegger's reading. Only the central claim concerns us here: That far from repudiating his metaphysical predecessors, Nietzsche – in a key respect (for Heidegger, *the* key respect) – remains very much in the Cartesian *subjectivist* grip. And seeing as modern metaphysics becomes first and foremost an an investigation into our *knowing*,[24] one would do well to look there. For it is precisely with regard to the problem of the possibility of knowledge, of *true* knowledge, where subjectivism with its theory of representations seems to bring us (and as we shall see, Nietzsche himself) to a halt – and thus where it must be dealt.

Before we get there, however, there's much that we must go through. First of all, we must deal with the fact that Nietzsche quite clearly took himself to be doing something very different. For one thing, he certainly did not think himself a part, let alone a follower of modern

24 That is, epistemology becomes "first philosophy".

metaphysics (or ancient, for that matter); he openly opposed himself to it. In fact, he didn't believe the Enlightenment project merely "incomplete" but, for the most part, fundamentally bankrupt. Built on wrong foundations. Glossing over this issue by saying that, in Nietzsche's eyes, those who came before simply didn't go far enough won't do. Instead we must inquire after these supposedly erroneous foundations themselves. What are they? What was Nietzsche truly against?

But here the answer is, I believe, quite obvious. For one finds it everywhere in Nietzsche – from the earliest works, such as *The Birth of Tragedy,* all the way to the mature pieces and unpublished notes. See, Heidegger only gives us half of the story. To appreciate the object of Nietzsche's critique in full, we would do better to turn to the other great "recollection" of western thought – Hegel's. Like Heidegger, Hegel sees the birth of modern philosophy in Descartes. But where Heidegger sees the decisive factor in *subjectivism*, for Hegel it is the affirmation of the "absoluteness" of self-determining Reason.[25] The supreme authority, as well as the self-grounding, all-embracing nature of *thought*.

> "Now we come (...) to what is properly the philosophy of the modern world, and we begin it with Descartes. Here, we may say, we are at home and, like the sailor after a long voyage, we can at last shout "Land ho." (...) The universal principle now is to hold fast to inwardness as such, to set dead externality and sheer authority aside and to look upon it as something not to be allowed. (...) What is deemed valid or what has to be acknowledged is thinking freely on its own account, and this can happen only through my thinking freely within myself; only in this way can it be authenticated for me."[26]

25 Incidentally, this Reason, when taken to its proper conclusions, is supposed to overcome the persistent spectre of subjectivism haunting Hegel's predecessors.

26 Hegel G.W.F., *Lectures on the History of Philosophy*, Volume III, University of California Press, 1990, pp. 131–2.

And it is here where Nietzsche is having none of it. He is is deeply convinced that the most profound error of all western though since Socrates is precisely such *rationalism* – the privileging of universal, impersonal reason and truth, at the expense of myth, instinct, and life. And it is with the modern period, of course, that he finds it at its most unfettered – seeing as the Enlightenment, with its main self-professed principle of accepting nothing but that which has been properly though-through (or "critiqued"), has has made the destruction of not only faith, but all myth and unreason its most explicit, universal task. – To build a rational society. How? By putting morality on rational grounds. How? By building on the autonomous, rational subject of Descartes.

This is why, despite its repudiation of supernaturalism, the Enlightenment can nevertheless seem as Nietzsche's greatest enemy. And it is also why Nietzsche turns to "psychology" – "the path to fundamental problems"[27] – to address this error at the roots. So, while Heidegger is right that Nietzsche remains *subjectivist* through and through, his *anti-rationalism* is just as essential.[28] In this respect, instead of not proceeding far enough, we should rather say that modernity went too far. Instead of recognizing the need for the *right kind* of myth and unreason, it became history's most powerful vehicle of nihilism.

However, theses such as these simply won't stand on their own. Nietzsche's complaints and exhortations cannot satisfy us if they stop at the pragmatic and prescriptive level – at his saying,

---

27 BGE 23.

28 In fact, they are merely two sides of one coin – mutually reinforcing one another. For if we cannot count on (objective) reason, what do we have left? To paraphrase Alasdair MacIntyre, if the rational and autonomous moral agent is exposed as nothing but a fiction, then the belief in morality (and thus morality itself) cannot be anything but a set of fictitious rationalizations. Rationalizations that conceal what? The fundamentally reality of the non-rational subjective *will*. Accordingly, in such a world, "my morality can only be what my will creates". "I myself must now bring into existence 'new tables of what is good.'" – The doctrine of the *Übermensch*. (See the chapter "Nietzsche or Aristotle?" in MacIntyre's *After Virtue*.)

"Truth is terrible; let us be superficial instead!" If Nietzsche's anti-rationalism and Hamletian pessimism – his belief that reason and truth are somehow inherently harmful and antithetical to life[29] – are to amount to something, he must first tell us what this *reason*, *truth*, and *life* actually are. Let us see what he has to say about about the first one.

29 In BT 7, where Nietzsche brings up Hamlet, he observes that his problem resides not in "thinking too much", but in thinking *too well*: "Dionysian man might be said to resemble Hamlet: both have looked deeply into the true nature of things, they have *gained knowledge* and are now loath to act (...) Knowledge kills action, for in order to act we require the veil of illusion."

II. Body and Soul

Nietzsche's overarching thesis concerning rationality and self-conscious though could be summed up by saying that, despite our everyday (perhaps even natural) beliefs to the contrary, it is in fact something highly *superficial*. And that in two main ways: First, in the sense that, when it comes to the question of human agency, reflection or rationality explain precious little of any consequence. "States of consciousness, any kind of belief, holding something to be true, for example," Nietzsche tells us, "are completely irrelevant and of the fifth rank when compared to the value of the instincts."[30] Second, in the sense that – *pace* Descartes who had to rely on God's perfection to ensure the veridicality of our ideas – our knowledge of the world (which includes self-knowledge) is necessarily not only incomplete, but essentially erroneous. For whatever it is that mediates between ourselves and the objective world, makes the object available to us, also inevitably twists it, distorts it, falsifies it – and thus turns it into something *other than* the sheer, objective fact – that is, a *subjective representation*.

The end of this work will be devoted to the critique of these two core theses by way of showing what they internally, logically entail. In this chapter, however, I begin by reconstructing the essential features of Nietzsche's own picture. First and foremost, I attempt to clarify some of the key "psychological" concepts and consider Nietzsche's position on the nature of the soul and body, mind and matter, and the need for their reconciliation in one coherent metaphysics. With this in hand, we can then, in the next chapter, delve a little deeper. Most importantly by taking up the central notion of Nietzsche's psychology, *the drive*, and by examining its role in the broader structure of agency and life.

---

30 A 39.

Now, the hermeneutic enterprise in here is rather tricky. For not only does Nietzsche use many ordinary "psychological" terms in not exactly ordinary ways, but seeing as the mind-body problem relates to just about every philosophical issue imaginable, trying to come up with a persuasive, consistent and coherent account of how Nietzsche deals with it would demand a work of its own. Still, metaphysics of agency can hardly avoid the problem. For that reason, I will outline Nietzsche's most important claims as well as offer possible ways of how to interpret them. However, because in many cases the sensible position is to accept that Nietzsche simply doesn't have a fully worked out account (as the more sensible interpreters recognize), instead of getting stuck on complex and possibly unsolvable exegetical puzzles, I will focus only on the issues most relevant to my purpose – and while doing so put the *philosophical* task first. Which is to say, rather than worrying about what Nietzsche "really meant", I will try to uncover what he *should have* meant.

~~~

In *The Gay Science*, we read that the problem of self-consciousness properly confronts us only when we realize just how easily we could do without it.[31] In this context, Nietzsche praises what he calls the "incomparable insight" of Leibniz: his thesis that self-consciousness, or *apperception*, is by no means a necessary condition of mental life and experience, not to mention the whole of it, but in fact only a small and contingent part.[32] And, as it turn out, one that's rather useless (if not outright pathological).[33] For Nietzsche bluntly declares that "we could think, feel, will,

31 GS 354.

32 GS 357.

33 See for example GS 11.
~~~

remember and also 'act' in every sense of the term"[34] even without this additional feature of seeing ourselves "reflected in a mirror".[35] Which is to say, while Nietzsche affirms we can be aware of (at least some of) our mental states and attitudes, this reflective dimension is not *constitutive* of them. After all, that is how things supposedly work in the rest of living nature.[36] And while we might be easily mislead into thinking otherwise, truly we are no exception: The vast majority of our inner life remains *unconscious.*[37] And in the rare cases where it *is* conscious,[38] as Mattia Riccardi aptly puts it, "this *extra bit* of our mental life does not play any role with regard to what we do".[39] Our consciousness thus turns out to be a mere surface phenomenon; an unessential by-product.

Remarkably, just a few lines bellow his assertion (that we could think, feel, will, remember and act even without becoming conscious of it), Nietzsche immediately contradicts himself when he asserts that consciousness *had to* develop because of our social natures, our profound individual helplessness and need for communication ("... the solitary and predatory man would

---

34 GS 354

35 *Ibid.* (Note that Nietzsche here commits himself to the notion of self-consciousness understood in terms of introspection – as if self-consciousness meant reflecting on one's mental states by way of another, distinct, "higher-order thought".)

36 *Ibid.*

37 See GS 333: "For the longest time, conscious thinking was considered to be the only thinking: only now does the truth dawn upon us that the greatest part of our mental activity goes on unconsciously and unfelt by us."

38 In most cases, Nietzsche uses the word "consciousness" (*Bewusstsein*) precisely in the sense of reflective (self-)consciousness – what he described as our ability to "see ourselves in a mirror": to be aware of, to know, what we feel, think, will and so on (rather than just unreflectively "do it"). Now this can make things rather awkward, as there are of course other important senses of the word. For example, Nietzsche clearly knows that non-human animals experience pain, or that they perceive the world around them. However, because they aren't reflectively conscious on his account, should we say these forms of awareness are "unconscious"? (For more on this topic, see Riccardi M., "Nietzsche's Pluralism about Consciousness," in *British Journal for the History of Philosophy*, 24(1), 2016, pp. 132–54; or for an even more thorough treatment, Part II of his *Nietzsche's Philosophical Psychology*.)

39 Riccardi M., *Nietzsche's Philosophical Psychology*, Oxford University Press, 2021, p. 130 (emphasis mine).

not have needed it").[40] How can something be inert and accidental but at the same same time necessary? Thus, contrary to the "epiphenomenalist" readings (which really only muddy the waters by introducing arbitrary requirements on what can and cannot count as a "real cause"), Nietzsche does seem to give consciousness a role – and therefore a use, a purpose. Even if somewhat inconsistently. As a matter of fact, just before presenting his hypothesis about the inseparability of consciousness and language, and their foundation in human sociality, Nietzsche poses the crucial question in these words: "What then is the point of consciousness when it is *in the main* superfluous?"[41] Indicating that, while it may be superfluous for the imaginary individual recluse, for socially habituated (that is, *actual*) human beings that is no longer the case.[42]

All that being said, it is safe to say that far from justifying panpsychism and the irreducibility of *the spiritual* in the manner of Leibniz, Nietzsche puts emphasis on the unreflective unconscious so that he can, in a way, do the exact opposite: reject its primacy, or significance in general. Accordingly, Nietzsche can be understood as subverting Leibniz's insight to substantiate his own claim about consciousness' superficiality – and to refer and bring this "surface" back to its ground – *the body*.

---

40 GS 354.

41 GS 354 (emphasis mine).

42 Riccardi, who claims that reflective consciousness is quite simply aetiologically inert, faces similar problems. In trying to make sense of the social aspect, he is forced to say that while an intention to act "does not causally contribute to the relevant action being produced, it turns out to be useful every time we are asked about the reasons for which we acted." (Riccardi M., *Nietzsche's Philosophical Psychology*, p. 153) It is simple to see that even if the first part were true, the latter – "giving reasons for which we acted" – is yet another action: one in which the intention is efficacious. Riccardi is aware of this (see p. 156); his solution, however, turns on the idea that while reflection does in fact possess some causal powers, these are merely "secondary" – themselves products of underlying "drives". But then, are the drives supposed to be uncaused? The inherently regressive mechanistic model is surely not the right one to make sense of agency.

> Behind your thoughts and feelings, my brother, stands a mighty ruler, an unknown sage – named Self. He dwells in your body, *he is your body*.[43]

And here, at least, Nietzsche's main purpose is clear. The soul must be naturalized, explained as a part of nature. And that to Nietzsche means: as something that could arise from the non-conscious. (Similarly to how life must be explained as emerging from inorganic nature.)[44] Accordingly, Nietzsche claims that reflective consciousness, as human "intellect" or "reason", is merely a "small reason",[45] subordinated, in the spirit of Pascal, Hobbes, or Hume, to affectivity and the unreflective, unconscious processes that go on "beneath the surface". That what we call reason or the conscious, reflective "I" is no more than "a tool" of the body – the "great reason" that governs it and follows its own *higher*, "ecological" (if not outright evolutionary), logic. As we learn in *Zarathustra,* we are bodies "through and through, and nothing else besides".[46] The body rules. And the soul itself is simply a word for something "about the body" (*am Lieb*).[47]

Now this can easily sound like a one-sided materialist reduction; one that effectively aims to eliminate the soul and the spiritual altogether. Yet this is not really Nietzsche's aim. For one, we do not know what, specifically, such "naturalization" ought to entail. The fact that the soul must be explained in a way science can work with is clearly not good enough. What we need to know is what explanations science allows and disallows. (Not to mention that science itself is a dynamic, continually transforming, ideally self-correcting collective project). But what's more, Nietzsche himself tells us that there is no need to reject the concept of the soul as such. After all,

---

43 Z "On the Despisers of the Body" (emphasis mine).

44 See for example GS 109.

45 Z "On the Despisers of the Body."

46 *Ibid*.

47 *Ibid*., (emphasis mine).

the soul, Ψυχή, is the defining subject matter of psychology, "the queen of the sciences". Nietzsche would be effectively arguing himself out of a job.[48] No, that is not the point. It is merely *a specific conception* of the soul that has no place in Science: namely, "the belief that regards the soul as something indestructible, eternal, indivisible, as a monad, as an atomon."[49] The real question is how the concept must be understood, what explanatory work it should be doing. Not whether it is "real". Accordingly, Nietzsche sets out to purify it – by repudiating all the moralistic, religious, superstitious baggage the "atomistic" notion of the disembodied "soul-thing" carries – and thus to rescue it as "one of the oldest and most venerable hypotheses"[50] we have.

So, what is this purified notion of the soul supposed to look like? For the reasons just mentioned, one of the ways into understanding Nietzsche's purpose here, the way he thinks (and wants us to think) about the soul, is to ask what he means by "psychology". Indirectly, that question already came up in the Introduction, and there it was conceived of (in an intentionally broad way) as the naturalistic investigation of agency. That is, an examination of what it means to be an agent, of the fundamental structure or logic of action. However, here it might be useful to view the matter from a slightly different angle. And for that, I will once again call onto Heidegger, who in his lectures on Nietzsche not only arrives at the same question but also provides an answer that we might use here to situate the soul, and "the psychological" in general, within Nietzsche's broader understanding of nature.

---

48 If it were to turn out that the notion of the soul is not (methodo-)logically necessary after all – that we can account for all phenomena without it – then psychology itself will prove redundant; of heuristic use at best.

49 BGE 12.

50 *Ibid*.

First, after asking about the meaning of psychology, Heidegger tells us what, for Nietzsche, psychology supposedly is not. (Although perhaps it would be better to say, what doesn't *exhaust* it.) That is: scientific-experimental research into mental processes; research into the higher life of intelligent mind; or "characterology", as the doctrine of various human types. Not even (philosophical) anthropology, as the inquiry into the essence of man, quite captures it; although here we are getting dangerously close. Only one clarification remains: for as Heidegger says, "Nietzsche's 'psychology' in no way restrict itself to man."[51] It is simply the question of "the psychical", that is, of *what is living*. This of course points us to nothing else than the very roots of the concept: The notion of the soul whose germs we find already in the Homeric epics and that gets eventually worked out into the Aristotelian notion of Ψυχή as the *principle of life*; or, which is the same thing, the "form" of a living being: that which makes it distinct *qua living*, as opposed to non-living (and dead). It shouldn't surprise us then that Nietzsche's psychology turns out to be inseparable from biology (and especially, physiology);[52] and that the soul or mind is thus always thought of as somehow "identical" – or *necessarily connected* – with the body.

To illustrate the philosophical force of this (Aristotelian) conception, let us briefly consider the problem of determining whether a given subject matter falls within biology. That is, determining whether the object in question ought to count as living. In order to do this, we have to look at the formal – which in this case means: *functional* – characteristics. For example, whether the entity exhibits, say, organisational closure, self-movement, metabolic exchange with the environment, or susceptibility to pain, disease, or in general, any kind of *purposiveness*.

---

51 Heidegger M., *Nietzsche*, Volume IV, p. 28.

52 In BGE Nietzsche even speaks of "physio-psychology" or "moral physiologists" (see sections 23 and 26).

Anything else, such as relying on specific "organic" material substrate, would be question-begging: Why is substrate A considered organic whereas substrate B isn't? Because it is functionally distinct; conductive to living self-organizing (living) beings. This way the soul can be said to be the living thing's "life-principle"; which in turn allows us to speak of *necessarily embodied* and mortal souls – as "animated matter" – without having to resort to any kind of vitalism, mystical forces, or theological assumptions; and without losing the organism *qua organism* altogether (something that is formally/functionally distinct from the inert, "inanimate" matter of the inorganic).[53]

In any case, what we have here is a perfectly neutral conception of the soul Nietzsche might accept. In fact, as long as he recognizes life as a distinct category – which he often does – he cannot really avoid accepting it (except nominally). But it is here that Heidegger warns us: When it comes to Nietzsche, we mustn't take "life" as a mere genus that includes humans, animals, and plants. Instead, it must be understood in the peculiar Nietzschean sense that determines all being as the active becoming, or "life", of the "will to power".[54] Now this is not the time for opening the can of worms that is the whole "will to power" doctrine; although we will get to that eventually too. The issue must be noted for another reason. Namely, because it is precisely at this point that Nietzsche's metaphysics threatens to turn into reductionism – a kind of substance monism: The idea that, ultimately, life and non-life aren't "really" distinct. That, in the end, life and agency *can be reduced* to the selfsame logic that governs the rest of nature; can be accounted for in the exact same terms, using identical principles – and thus dissolved.[55]

---

53 In BGE 12 Nietzsche speaks of "those clumsy naturalists" who "can hardly touch 'the soul' without losing it".

54 Heidegger M., *Nietzsche*, Volume IV, p. 28.

55 The argument is simple: If the organic and inorganic are indeed at bottom one, and the very distinction between them is a "hasty prejudice" (WP 655), then the former's essence (whatever it may be: mechanism, the

That something like this is the case with Nietzsche is quite evident from a number of passages (though mostly unpublished ones).[56] But of course we are under no obligation to follow Nietzsche here. In fact, in the majority of his writing he clearly treats the distinction as very much real and important (even if he wants say that, yes, at the end of the day even this might be done away with; that the distinction between the living and the non-living, purposive and non-purposive, is a distinction that is only *subjectively* necessary, a regulative idea, or something like that). Thus, for now, we can safely stay at this level of explanation, the level where living, purposive organisms *are* (or at least *seem*) formally and functionally distinct from non-purposive inanimate objects;[57] and where the talk of agents and agency doesn't come out empty, as a mere heuristic. Then, eventually, we shall see whether we might indeed go "deeper" and entitle ourselves to any such reductive move.[58]

But let us now return to what Nietzsche himself has to say about *die Seele*. So far, we have seen him reject the empty atomistic "soul-thing", and instead affirm that the soul must be thought of as embodied: inseparable from the body, even, in some respect, identical with it. But he has other suggestions on how to conceive of our psyches properly. Right after his dismissal of

---

will to power, etc.) will also be the essence of the latter. Would this effectively make Nietzsche a kind of panpsychist, an heir of Leibniz or Spinoza? Nietzsche himself would undoubtedly object to these labels on the grounds that they imply the exact opposite of his own position: Panpsychism says that everything is alive, and thus stands things on their head; whereas Nietzsche emphasises that what we normally call "the organic" is but a recent development, something that emerged accidentally and is, in comparison, derivative and rare. As such, as he says in GS 109: "the living is merely a species of the dead." However, all of this proves rather pointless. Whether we reduce the organic to the inorganic or *vice versa*, both poles lose all meaning. The very concept of life becomes superfluous – merely a different *name* for being. When everything is alive, *nothing is*. This is the g problem of all substance monism: it is constitutively unable and ground distinctions, and thus get off the ground.

56 See for example WP 635, 655, 689; KSA 9:11[70]; but also BGE 22, 36, and perhaps most importantly, GS 109.

57 See WP 521.

58 The question of the irreducibility of life and agency is the main topic of the last chapter.

atomism and dualism, he claims not only that the souls must now become "mortal" (which shouldn't surprise us at this point, given that it seems entailed in the thesis about essential embodiment), but also introduces the notions of the "soul of subject-multiplicity" and the "soul as social structure of drives and affects".[59]

What are we to make of these? One thing that seems immediately clear is that the Nietzschean soul will be something composite or complex. Again, we may think of the classical theories. Even according to Plato, the soul is a tripartite "social structure" of sorts (although of course, as we learn in the *Phaedo*, it is also immortal). One in which the three basic elements (the rational, spirited, and appetitive) exist in a kind of hierarchical order can come to predominate in varying degrees and thus determine, to put it crudely, what sort of person one is. That is, what things one values and pursues. Now, according to Plato, this works both ways. By learning to value and pursue the right things, we can bring harmony into our souls – for there is a "natural order" to how the soul should be – and thus become different persons.[60] With Nietzsche, things are ... more complicated. For one, he certainly does not think one achieves any kind of harmony by prioritizing reason or deliberative "self-control" over one's appetites and instincts.[61] And the notion of a "natural" hierarchy would also need to be seriously qualified (that is, relativized).

---

59 BGE 12.

60 As the *Republic* and *Phaedrus* make clear.

61 See especially the last few sections of "The Problem of Socrates" in *Twilight*. For example, from section 11: "To *have to* fights the instincts – that is the formula for decadence: so long as life is *advancing*, happiness is equal to instinct." Of course, then there is Nietzsche's analysis of "Socraticism", which comprises the most significant part of *The Birth of the Tragedy*. Or the crucial section 191 of *Beyond Good and Evil*, where Nietzsche tries to expose Platonic reason as motivationally inert and necessarily subordinated to instinct – and thus all moral judgements as articles of "irrational faith".

Nevertheless, the basic idea that we should understand the soul as a hierarchical, almost "political" structure that determines one's normative commitments is very much at work here.[62]

The basic rationale behind conceiving of our "inner" in terms of a composite, hierarchical structure is not hard to see. It seems like the simplest way to account for the ubiquitous experience of psychological *conflict*:[63] the fact that our motives, values, desires, and beliefs are not just multiple, but often come to clash with one another. In fact, only a few pages after discussing the concept of the soul, Nietzsche turns to examining the logic and the phenomenology of *willing*. And there he writes: "A person who wills – commands something inside himself that obeys, or that he believes to obey".[64] And that is supposedly made possible because "our body is a society composed out of many souls."[65]

Now this might seem to complicate things. Taken together with the aforementioned definition of the soul as a "social structure of drives and affects", we might be lead to think that we have a body that is made up of "souls", which are in turn comprised of "drives and affects". That seems not only extremely bizarre, it also doesn't get us anywhere. Yet there is a simple way out. All we need to do is take Zarathustra's words from above seriously: The body and the soul are not just connected, they are *one* – united in what Nietzsche calls the Self; or, *der Lieb*: a *living* and "creative" body – "a great reason, a multiplicity with one purpose, a war and a peace, a herd and a herdsman."[66] Thus, when Nietzsche speak of the body that consists of "many souls" he is

---

62 See mainly BGE 268.

63 On the other hand, it also introduces some serious problems. One, there is the danger of *reifying* the soul again. And two, how is it that all this compositeness results in the final, seemingly irreducible "apperceptive unity" – the fact that all these parts, or their effects, are experienced *as mine*?

64 BGE 19.

65 *Ibid*.

66 Z "On the Despisers of the Body."

not referring to anything other than those very unconscious "drives"[67] that constitute *the* soul – and therefore the body – and therefore our selves as such.[68] This is Nietzsche's conception of the body/soul *in nuce*. A unity composed of and governed by an unobservable structure of a variety of interrelated elements: a "social structure", "a subject multiplicity", a "rank order"[69] of the *drives* – the primary constituents of our very "being".[70]

As such, our next task becomes clear. We must turn our attention towards these drives themselves. Why ever conceive of them as "souls"? How do the individual drives relate to each other? Or to the organism as such? And what is their relationship to self-consciousness and reason?

---

67 I omit the "affects" here because they are essentially included in the drives – they are the "ends" the individual drives aim at, and thus what results (*qua* a determinate phenomenal state, a specific qualitative sensation, and affect, a pathos) when that aim is achieved, when the drive can "discharge itself".

68 For a similar reading see Riccardi M., "Nietzsche on the Embodiment of Mind and Self," in Constâncio J. & Branco M.J.M. & Ryan B. (eds.), *Nietzsche and the Problem of Subjectivity*, De Gruyter, 2015.

69 BGE 268.

70 D 119. (Nietzsche's point is that each of is a totality of specific drives, and these determine what kind of person one ultimately is. That is, they constitute one's being, but this "being" is specific to every individual. The word he uses is *Wissen*, which could be just as easily translated as "essence", "nature", but also as "*character*".)

III. The Drives

In one particularly interesting note, Nietzsche writes: "Each 'drive' is the drive towards 'something good', as seen from a particular perspective; there's valuation in it."[71] This, considered along the fact that the drives are the fundamental constituents of our being, shows us that our relation to the world is always value-laden and evaluative. But it also gives us a rationale for why Nietzsche might conceived of the drives as "souls". Namely, as Riccardi notes, Nietzsche might be emphasizing their *mental*, rather than physical, aspect.[72] Specifically, the fact that they bear a proto-*intentional* form – their being "aimed at something", something that is at the same time interpreted or valued as "good"; that is, a goal, an end, a *purpose*. (As Nietzsche has it, this good is the discharge of the drive and the *affect* that accompanies it, but we don't have to worry about that now.)

This intentionality or "directedness" of the drives, it would seem, is also how we can know, at least in some limited way, what drives there are, and even that they exist in the first place. For after all, they are unconscious; we have no direct, introspective access to them.[73] Which is also why, according to Nietzsche, they have gone hitherto virtually unnoticed. As such, it seems only logical that the only way to discern them is by empirical, *a posteriori* inference based on our noticing of certain regularities or patterns of activity the organism under observation exhibits: its specific functional aims, the goods it pursues, and in the case of self-knowledge, presumably also the specific phenomenal states (the affects) linked to them. For instance, one may identify the

71 KSA 11:167.

72 Riccardi M., "Nietzsche on the Embodiment of Mind and Self," in Constâncio J. & Branco M.J.M. & Ryan B. (eds.), *Nietzsche and the Problem of Subjectivity*, p. 538.

73 See D 119.

drive to nourishment by the activity of searching for food accompanied by the unpleasant feeling of hunger (a certain qualitative, phenomenal "lack"), as well as the pleasant feeling of satiation that arrives upon hunger's satisfaction (a certain qualitative "fullness"). Accordingly, the drives thus *motivate* and direct the organism towards certain ends or activities – ones that usually require pursuing and preferring particular *objects*. In the case of hunger, the objects one takes as the possible source of nourishment, the *satisfaction* of that hunger – the *aim* of the drive, its gratification or "discharge" in this or that particular object.[74]

Moreover, the drives also affect how we *perceive* and *interpret* our environment. Because of their motivating directedness, we embody an *evaluative perspective* on the world, in light of which the world appears as "practically significant" to us; as filled with objects we take *as* gratifying, *as* obstacles, *as* neutral, etc. In the words of Katsafanas, the drives "determine which features of the agent's environment will be salient" as well as "the way in which the agent will conceptualize aspects of the environment".[75] Crucially, this means that even perceptual and theoretical activity must be understood as subjectively purposive, and thus as species of *practical* activity. As forms of valenced responsiveness to external conditions that show up as "good-for" and "bad-for", instead of some passive, disinterested (perceptual or cognitive) "registering" of neutral, objective "facts".[76]

But now, what about these drives makes them the primary constituent of one's being? All I've said so far could just as easily apply to regular desires. Nietzsche's answer is that the drives are

---

74 See D 119.

75 Katsafanas P., *The Nietzschean Self: Moral Psychology, Agency, and the Unconscious*, Oxford University Press, 2016, p. 110.

76 As Nietzsche writes in WP 505: "There is no doubt that all our sense-perceptions are wholly permeated by valuations (useful or harmful — consequently, pleasant or painful). (...) Insects likewise react in different ways to different colours: some like this shade, the others that."

quite simply more "basic" and "constant". That is, rather than specific, discrete desires (although it might already be a mistake to think of desires in this way, as I will discuss bellow) they are more akin to *needs*. As such, they cannot be "eliminated" by attainting their given aims; they are not static finite "goals" that can be achieved "for good". Each drive may be temporarily subdued, to be sure, but it *will* come knocking again. Accordingly, the best way to conceive of them might be as of *behavioural dispositions*: the stable elements that constitute one's fundamental psychological makeup and ground the other, more immediate states.

~~~

All of this seems to confirm that the Nietzschean drives – as the most basic psychological explanatory posits – are indeed by no means simply identified with, or reduced to, mere physiological states. Or more precisely, given our discussion about the body and soul above, they *are* physiological states, necessarily biologically realized, but that is not *all* they are. The psychological component is also essential. Even Leiter and Riccardi, who both believe that Nietzsche's naturalism ultimately amounts to giving a non-normative causal account,[77] agree that Nietzsche "was fairly explicit in rejecting any type-identity of mental and physical states of the person"[78] so that "psychological explanations are autonomous, not reducible to merely biological and physical ones".[79] According to Riccardi, this yields a kind of "non-reductive physicalism".[80] The question is, of course, what that is, exactly. If it meant only that we simply cannot transcend the material aspect of existence in favour of existing in some higher spirituality,

77 See Leiter B., *Moral Psychology with Nietzsche*, Oxford University Press, 2019, p. 100; Riccardi M., *Nietzsche's Philosophical Psychology*, p. 65.

78 Leiter B., *Moral Psychology with Nietzsche*, p. 69.

79 Riccardi M., *Nietzsche's Philosophical Psychology*, p. 55.

80 *Ibid*.
~~~

then all would be well and good. But rejecting dualism is often easier said that done. What usually happens instead is that one ends up with having to juggle two autonomous kinds of "things": mental and physical. And that won't do. But we already have a framework to accommodate this. The functionalist, hylomorphic one of course. For there, instead of the dualistic autonomy of physical and mental substances, what we get is the constitutive inter-dependence of "matter" and "form"; such that both sides only "make sense" with respect to one another, become intelligible through one another, as two internally relating moments of a unified whole: the "thing" (or substance) we are trying to explain. – The body, to be a (living) body, must be ensouled; "informed" in such a way that its matter exhibits the characteristic function of living bodies. And equally, the soul is nothing on its own, if it isn't embodied or "enmattered".

Now Nietzsche himself never explicitly affirms hylomorphism. In fact, he considers both "matter" and "form" (and even more so, "substance") nothing but anthropomorphic constructs that we must project onto the formless reality of chaotic sensations to make existence bearable for us.[81] Yet, given that he himself must (and does) project it so, once again, we can simply stay at this level of analysis for now. And we need to, for this issue is far from auxiliary – primarily because of yet another (albeit closely connected) reason. For we are now getting dangerously close to one of the most formidable challenges the Nietzschean soul and drives have to face. The challenge of the so called *homuncular fallacy*: the attempt to account for the behaviour of organisms or agents by breaking them down into "micro-agents" that posses simpler yet still irreducibly "agential" properties (so that the resultant "explanation" turns out to be question-

81 See for WP 523, 569; GS 109.

begging or mereologically fallacious.) And indeed there are many passages where Nietzsche seems to be doing just that. Just consider: He tells us that the drives, among other things, entail "perspectival valuation",[82] "interpret the world",[83] or that "each one has its perspective that it would like to compel all the other drives to accept as a norm."[84] While this sure packs into the concept of the drive some of the essential marks of the mental, such as subjectivity, intentionality, and normativity (which might reinforce the idea that Nietzsche's psychology does in fact part with crudely materialistic explanations), it does so using language that should give us pause.

There are many ways one could go here. Some take Nietzsche's language literally and think him a hopeless homuncularist, a mental partitionist through and through.[85] As most of them realize, that would be fatal. Yet despite all the overt homuncular and mentalistic/intentionalistic talk,[86] Nietzsche does give us signs that might point elsewhere. The most explicit one of which, I believe, comes when – after talking about drives that "desire" things, "exercise" and "discharge" their strength, and engage in all sort of other agential activities – Nietzsche interjects with an uncharacteristic: "these are all metaphors".[87] That is undoubtedly an important acknowledgment on Nietzsche's part. The question is: what now? For it might very well just make us nervous that he simply doesn't have a clear enough concept at all.

---

82 See KSA 12:1[58] (but also HH I.32 or BGE 6).

83 WP 481.

84 *Ibid*.

85 See Gardner S., "Nietzsche and Freud: the 'I' and Its Drives," in Constâncio J. & Branco M.J.M. & Ryan B. (eds.), *Nietzsche and the Problem of Subjectivity*, De Gruyter, 2015; or Poellner P., *Nietzsche and Metaphysics*, Oxford University Press, 1995, p. 215–20.

86 Cf. WP 490.

87 D 109.

To answer some of these worries, it has become standard to turn to the *dispositional* reading of the drives put forth by Richardson (first in a more generic form in *Nietzsche's System*, and later in a more evolutionary garb in *Nietzsche's New Darwinism*). On this account, saying that a drive aims at certain ends or views the world in a certain way does not mean the drive is somehow personal or conscious.[88] The drive is merely a behavioural "disposition that was selected for a certain result;" and where "this result is its individuating goal, which explains its presence and its character."[89] On the surface, so far, this seems quite plausible. Consider, for example, a migrating bird: It has never seen winter and likely never will. Yet, somehow, it almost invariably flies to the right place at the right time. How does it know? Well, his drives seem to "know" (although today we would probably say genes).

Among Nietzsche scholars today, there is almost universal acceptance that this (or something close to it) is in fact what the Nietzschean drive is.[90] Katsafanas in particular has pursued this line of thought in great detail.[91] Yet we must tread very carefully here. Homuncularism and the "pathetic fallacy" are no strangers to many a living philosopher (or biologist, for that matter), and it is far from clear that this strategy is how we avoid or overcome them. Indeed, as far as I can see, everything depends on how exactly we understand Katsafanas' claim that "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

88 Richardson J., *Nietzsche's System*, Oxford University Press, 1996, p. 38.

89 Richardson J., *Nietzsche's New Darwinism*, Oxford University Press, 2004, p. 39.

90 In some shape or form, Katsafanas, Leiter, Riccardi, Clark, Dudrick, or Janaway (among others) have come out in favour of it.

91 See Katsafanas P., "Nietzsche's Philosophical Psychology," in Richardson J. & Gemes K. (eds.), *Oxford Handbook of Nietzsche*, Oxford University Press, 2013, pp. 727-755; or the "Drives" chapter in Katsafanas P., *The Nietzschean Self*.

reason, evaluate, and interpret".[92] For it is at this point that homuncularism meets (or threatens to turn into) reductionism. To illustrate the problem, consider that while we have no problem saying that the heart, as part of the whole organism, can in fact pump blood (which is a purely physiological process), we do not usually speak of eyes that can see – provided that by sight we mean something that contains intentional activity or normative *judgement*. Rather, it is the organism itself that sees; and it can do so because it has (physiologically) well functioning eyes.

A highly instructive case in this debate is Riccardi; seeing as he has recently attempted to use the dispositional picture to ground his deflationary reductionism about the normative sphere, as the following makes clear:

> On the one hand, on my reading the drives do not need to be illegitimately equipped with personal-level capacities like sensation and consciousness. On the other hand, I do not assume that Nietzsche's normative vocabulary – his talk of drives 'dominating', 'commanding', 'obeying', etc. – is irreducibly so. For my strategy consists precisely in showing that, for all Nietzsche says, the states and processes he describes by appeal to such normative terms can be further analysed into simpler states and processes describable in non-normative, dispositional terms.[93]

Riccardi believes his deflationary account "frees Nietzsche's psychology of drives from any charge of fallacious homuncularism".[94] Yet that claim rests on his accepting that there exists a viable, non-fallacious form of the method: a strategy wherein we still proceed by breaking the agent down into smaller micro-agents, but carefully, such that we do not end up with the various logical vacuities.[95] In other words, Riccardi also thinks that the drives, on their own, cannot be

92 Katsafanas P., *The Nietzschean Self*, p. 97.

93 Riccardi M., *Nietzsche's Philosophical Psychology*, p. 65.

94 *Ibid*., p. 48.

95 *Ibid*., p. 49–50.

said to reason, interpret, and so on; and that it is only when considered as parts of the organism that such language starts making sense. First we must address the problem of the reduction: the idea that we can get rid of the normativity of the agent (or soul) as such, by examining its parts (the drives). The decisive question is this: Is this strategy supposed to give us a genuine – more fundamental, scientific, perhaps even more "real" – *explanation* of the agent (rather than the individual parts in question)? Riccardi, following Dennett, thinks that is indeed so: The *normative soul* of the agent, and thus the agent as a whole, as such, is – in the final line of analysis – supposedly reducible to its constitutive *non-normative* components: the *dispositionally-causal* drives.[96]

But is this reductionism not simply *atomism* – the one doctrine Nietzsche tells us has no place in any sphere of science? For what atomism believes (and has believed ever since the ancient "atomists") is that we can – and must – explain things, account for what they really are "in themselves", precisely by breaking them down into their constitutive (material) *parts* and the causal processes between them (the atoms and their collisions, "swerves", and so on). This is exactly what Riccardi suggests.[97] And it is exactly what the functionalist critique exposes as

96 "What the normative reading aims at establishing is not merely the – to my eyes, relatively uncontroversial – claim that there is a normative side to Nietzsche's characterization of the soul and of the drives (and affects) constituting it [but] that the normative dimension which is part of that characterization is irreducibly so." (Riccardi M., Nietzsche's Philosophical Psychology, p. 56) And: "On the normative reading defended by Clark and Dudrick, the relations among the drives Nietzsche describes in agential terms cannot be analysed into merely causal interactions." (*Ibid*., p. 65) Whereas Riccardi's "strategy consists precisely in showing that, for all Nietzsche says, the states and processes he describes by appeal to such normative terms can be further analysed into simpler states and processes describable in non-normative, dispositional terms." (*Ibid*.)

97 Riccardi insists on the irreducibility of the mental (his "non-reductive physicalism" and the "autonomy" of the psychological), but what can this mean now? Are we to preserve the mental domain while abolishing the intentional and normative? Is the mental/physical just some convenient way to partition our explanations or their object? If so, what makes them different? How are the "micro-agents", if they aren't intentional, not just physical *organs*? – The answer, of course, is that "non-reductive physicalism" is really just material reductionism with a rhetorical flourish added on top.

woefully insufficient. Of course, nobody disputes that there are great many things to be found by searching the inmost *parts* of organisms' bodies. The soul, mind, "selfhood", or consciousness, however, are most certainly not among them. *Not* because the soul is some immaterial, imperceptible, mystical *res cogitans* or monad. But because it is nothing *materially distinct* from that body itself. Should that mean we cannot distinguish[98] between *minded*, intentional *actions* and *mindless*, mechanical *events*? Of course not. It means only that to account for that (absolutely crucial) distinction, we are going to need more than just crude matter (whatever that is) and efficient causality.[99]

In fact, when it comes to living organisms, even when trying to understand their physical organs the atomistic endeavour proves completely hopeless. Why? Because *function* can never be understood or accounted for as a mere emergent property of "matter". Rather, it is the principle of its organization – its *form*. Thus, material reductionism gets the explanatory priority the wrong way around: In accounting for the relation between matter and form, between material processes and purposive activity, it thinks the former more fundamental. This way, however, it makes emergence purely one-directional; the relationship between causal processes and purposes purely accidental – and thus turns all function (such as purposive intentionality) into nothing

98 Logically; or, *formally*.

99 Even efficient causality, as understood in the context of Aristotle's famous four-fold nexus, differs quite significantly from causality as regularly understood today. True, both principles are ways to explain "what is" by showing what *antecedent* brought it about. But "modern", mechanical causality is a "physicalist" causality: one that concerns exchanges of energy through antecedent forces that work upon some mass (in fact, *that's what the cause is*). Aristotle's causality, on the other hand, is a *logical* one. That is, the αἰτία are not so much "causes" as necessary explanatory principles – mutually implicated rationales, or logical relations; the constitutive "grammar" of predication without which what exists, insofar as it exists, *would simply not obtain* (would not be explainable; couldn't be accounted for). For that reason, it would also be a gross distortion to think of, say, the final cause – modern science's *bête noire* – as if it was some magical force pulling onto (and extrinsic to) matter (itself understood in diametrically opposed way) from "up above".

but an observer-relative "fiction", non-essential to the entity *in itself*.[100] But – and this is the main point – matter can be the matter it is only in virtue of its form; only on the basis of a function. Without it, there is no determinate entity to speak of. Accordingly, on the functionalist account, part and whole specify each other in a reciprocal fashion (and it the form/function that has explanatory primacy).[101] The reduction of a heart to its material parts won't tell us what makes heart a heart. Instead, by subtracting the functional dimension – the "final cause" of pumping blood – we lose the heart (as the phenomenon *to be explained*) completely.

For the same reason, one also has to be careful when trying to account for function via natural selection (as Richardson suggest with relation to the drives). Why? Because the key evolutionary notion of *adaptation* – the driver of selection – is intelligible only in relation to some standard of "good" and "bad" (or "fit" and "unfit", if we want). But notice: this *presupposes* functionally organized systems. Normative, internally purposive beings – the very things we are trying to explain. Indeed without internally purposive beings (beings that "care about their own being"), there would be no evolution to speak of. For evolution the governing immanent dynamic of *life* itself; that is, of the *reciprocal* interactions of purposive organisms with their contingent environment. Not some reified external force – "Mother Nature" – externally *designing* things, as only an agent could.

---

100 That is, the entities only *seem* normative and intentional *to us* who are trying to explain and predict their behaviour. The "intentional stance" is merely an "as if"; a *regulative* idea, rather than a *constitutive* one.

101 Kronz and Tiehen correctly observe that "it does not make sense to talk about reducing an emergent whole to its parts, since the parts are in some sense constructs of our characterization of the whole." (Kronz F.M. & Tiehen J.T., "Emergence and Quantum Mechanics," in *Philosophy of Science*, 69(2), June 2002, pp. 324-47)

Anyway, we can now return to the decisive question. The question about the status of the drives, first with regard to the organism as a whole, second with regard to the question of the physical and mental. Concerning the former, we now see that the drives are simply unintelligible on their own; and that they must be understood as subordinated to the function or form of self-maintaining organism, whose every part is at the same time both end and means.[102] In fact, without this purposive unity, even the concept of *embodiment* just becomes meaningless. Now this flows directly from the functionalist critique. But there is still the second question– the claim that drives, as embodied, "considered as part of a whole organism" (rather than the purposive unity of the organism as such), *can* interpret, evaluate, reason and whatnot. What should we say here? – Not only does this explain absolutely nothing; it makes things worse. For one, and crucially, how do we now account for the intentionality as well as the very unity of the organism itself? For "selfhood", the "I", the "synthetic unity of apperception"? How could there ever be an agent *composed of* drives? How could there be experience (*Erfahrung*) – distinguishing between different experiences (*Erlebnis*) or states (say, of the affective or other responses elicited by the various drives) – if there is no *I* doing the distinguishing? If the drives were intentional agents, there would be as many unrelated "selves" as states and so no alteration to speak of – making the very notions of experience and the self unintelligible.[103]

The overall result we have reached is thus the sensible one. The drives cannot be some occult little imps perpetually deceiving us from somewhere within our bodies. They cannot be

---

102 The purpose of the heart is to pump blood to the vital organs; in turn, these organs keep the heart capable of pumping; and the overall purpose – the ultimate rationale, αἰτία or "cause", of the whole thing – is the continual functioning, or maintenance, of the organism itself.

103 Overall, this position is really nothing else than the "empiricist" or Humean theory of the Self, and thus suffers from the very same – fatal – problems. I will say more on the matter in the last chapter.

intentional at all because intentionality makes sense only with respect to unified purposive agents capable of lived experience. Agents that can *act*, that care about themselves, can fail, and ultimately, die. And that returns us to the main lesson of hylomorphism again: Intentionality, consciousness, or selfhood cannot possibly be some *properties* of pre-existing substance. They can only be substantial *forms*. That which is unifying (and embodied in) such experiencing, acting, or believing; the very self-constituting *activity* – the ἐνέργεια – of purposive living. Not some-*thing* separate that accompanies it. Consequently, seeing as the drives, even as necessarily embodied, cannot really *do* anything (not "interpret", "evaluate", "desire", "reason", or "adopt perspectives"), we can now also appreciate why Nietzsche – in speaking of the drives in isolation, (identified on the basis of our *specific* aims, i.e. some signature patterns of activity) – cannot avoid turning to abstractions that inevitably call for metaphorical (agential and homuncular) language.

IV. Life as Will to Power

However homuncular or dualistic we want our Nietzsche to be (and who knows, perhaps he really was a "genocentric" evolutionist long *avant la lettre*), I now leave the drives and their mutual interactions behind. In the last chapter it became clear that what we need to focus on – in investigating the logic of agency as such – is rather the concept of the organism itself. And so this is where I turn now. As it happens, however, a great deal of the work has already been done. For seeing as the notion of the drive was, for Nietzsche, supposed to represent the most elementary "logic of living", the basic dynamic of intentional, purposive striving towards its satisfaction or "fulfilment", we can simply put the metaphorical language aside, take the notion of embodiment and purposive unity seriously, and see what sort of being emerges.

We may begin by recalling Heidegger's words from Chapter II. There we saw that Nietzsche's psychology is really nothing other than that which deals with the "psychical", that is, with the living; and that life for Nietzsche signifies some specific "way" of being – a mode or form of "being a being". What form? The one that determines that being as the "will to power". Now, granted, according to Heidegger's Nietzsche – the last great metaphysician of the West for whom the will to power represents the answer to the basic metaphysical riddle about the essential character of *all beings as such* – this "will" can hardly help us determine the essence of the living *qua living* (that is, as distinct from the non-living). On this "inclusive" reading, the will to power is a monist, ontological and cosmological "theory of everything" *à la* Schopenhauer; and "life" is thus really little more than a name for Being. This allows Heidegger to conclude (not entirely without warrant, it should be said) that insofar as the will to power "constitutes the basic character of all beings, and inasmuch as the truth of the whole of beings as

such is called 'metaphysics', Nietzsche's 'psychology' is simply coterminous with metaphysics."[104] But of course, once again, there is no need – or reason, really – to follow (Heidegger's) Nietzsche there.[105] We can be perfectly satisfied with narrowing our scope down and treating life as something distinct from non-life. Which, as a matter of fact, is exactly what Nietzsche does most of the time.[106]

However, if we do that, the will to power, trying as it does to capture the peculiar character of this form of being, amounts to nothing other than Nietzsche's account of subjectivity: of what it means to be an embodied subject, an agent, or plainly: a living being. Yet as I already remarked, all this is implicitly present in the nature of the drive. All we have to do is flesh out the details – and perhaps expand on some of the more significant features. Let us therefore begin with the question of what the elementary intentional structure of the drive, the drive *towards* "something good" – now interpreted as the internal dynamic of living – entails; and what thus necessarily follows from its concept.

---

104 Heidegger M., *Nietzsche*, Volume IV, p. 28

105 In footnote 55 above, I provide the basic reasoning for why we should actively resist all panpyschist theories (including Nietzsche's "anti-panpsychism" where *nothing* is alive); and more broadly, all kinds of Parmenidean/Spinozist "substance monism" (of which Nietzsche's late theory is undoubtedly a species, although I don't have the room here to actually make that case).

106 For what it's worth, Heidegger's Nietzsche is rather hard to find in Nietzsche's actual texts, the one exception being the late unpublished notebooks. (Of course Heidegger thought that's precisely where the "true Nietzsche" is hidden – a dubious belief at best.) Aside from the *Nachlass*, there are only two mentions of the will to power being this "exceptionless" and "absolute" essence of the world (namely, BGE 22, 36). Although in GS 109, Nietzsche is quite explicit that the distinction between the organic and inorganic is our own "imposition" on things. That being said, most mentions of the "will to power" simply treat it as the fundamental principle of living (i.e. organic) things. For example, in the famous chapter "On Self-Overcoming", Zarathustra describes it, in quite the romantic spirit, as the "inexhaustible procreative life-force" and says: "Wherever I found a living thing, there I found the will to power." (For some other mentions along these lines see: GS 349; BGE 13, 259; A 6; GM II.12.)

First, for there to be a "drive" or striving *towards* anything, life presupposes some sort of a *need* – a "negativity" that manifest itself as self-referential – or *felt* – deficiency or lack. For that striving is nothing other than life's primary "impulse" to close this gap, to satisfy the need – the *conatus* for *fulfilment*: the internal final cause or purpose of life.[107] And yet, because this negativity is not merely negative, not merely a restriction, but positive as well – because it has the constitutive or enabling role of opening up the "space of purposes" (the space for there to be any striving and life in the first place) – the organism can only exist in a state of inextricable dependency on what is other than itself, its world or environment. This has some important consequences; the most crucial of which is that such fulfilment simply *cannot* mean freeing oneself from this negativity altogether, as if to achieve some pure state of absolute autarky. This is surely *the* conflation to be avoided; the most pernicious mistake one can make here (politically and otherwise).[108]

In other words, this logic of striving should most decidedly *not* be taken as implying some ultimate metaphysical futility of life (like, say, the "unhappy consciousness" of Sartre). Although striving certainly implies negativity and rightfully carries some negative connotations, nothing here says that our ends in life are *a priori* unattainable. On the contrary, if we follow our logic closely we learn that life's negativity – the ability to suffer (experience loss in the most expansive sense of the word), and ultimately, the possibility of life's actual negation (its susceptibility to

---

107 In the first book of the *Nicomachean Ethics*, Aristotle famously deduces that life must exhibit such a structure; that is, that εὐδαιμονία must be the absolute purpose or highest "good": a purpose unto itself, the only "thing" pursued for its own sake. I will attempt to make this logic, along with the real meaning of such absolute purpose, more precise in what follows (this and the next chapter).

108 Unfortunately, Nietzsche – with his language of absolute, unflinching affirmation – sometimes does fall into this trap. The famous, putatively affirmative ideals of *amor fati* and "eternal recurrence" are, I believe, an unfortunate recourse to the "logic of transcendence" Nietzsche otherwise strove to overcome; and the ultimate consequence of this dangerous conflation. (To illustrate the difference, consider that to fully affirm *mortal life* doesn't mean affirming or embracing death. No, death, in itself, is still very much negative; but it is necessary for there to be life in the first place. In other word, affirmation of finite life must mean a struggle against death.)

death: the loss of all purposes) – is not a restriction but an *a priori* constitutive *condition* of having any kind of purpose in the first place. And that entails the possibility of *both* failing and succeeding at maintaining it. Thus, far from affirming the impossibility of fulfilment – as if it existed at an infinite distance from us, in an unreachable *beyond* – we see merely that fulfilment cannot be a "goal" that could be accomplished once and for all, something that could be *finished*. No, it truly *is* an *infinite* end. But not because it is unreachable in this life. It is infinite because of the simple fact that it must always be *sustained*. For even at life's most fulfilled, the *possibility* of failure – as the minimal condition of purposive life at all[109] – must always remain. But that clearly doesn't mean success can't be *actual.*

What we come to see is that the *dependence* of any life on its "other" (on non-life and usually other life) cannot be overcome because independence isn't its mere negation, *non*-dependence as its absolute opposite, but rather a more complete, "higher" form of its manifestation.[110] Accordingly, fulfilment is the transcendental end of all desire; not the transcendence of desire. To wish for such transcendence, for freedom from life's negativity *tout court* (i.e. to close the gap once and for all) – is nothing but to wish for death.[111] For the real opposition, the "tension of life", isn't between the organism and some external standard. It cuts through the organism's *own self* – as the difference between what it is and what it strives to be.

---

109 Or perhaps we should rather say the *maximal* condition, as the minimal condition is of course that of simply *staying alive* (staving off death). Life happens between them.

110 A "sublation" (*Aufhebung*) to use the notorious term of art of Hegel's.

111 What Freud's famous "death drive" should have meant.

It is precisely *this* dynamic, negating self-relation of striving and "overcoming", this "*Ur*-form to which all drives can be traced back,"[112] that Nietzsche likes to call *der Wille zur Macht* – the will to power – *which is just the will to live.*[113]

On this account, being a living being thus turns out to be nothing else than the activity to be a living being, to be itself. This means that life must be understood *as its own purpose* (hence *internal* purposiveness); or better, that living beings must be understood as "purposes unto themselves": self-maintaining "systems"; entities that are "for themselves" – which in this context means: beings that act *for the sake of themselves* (for the sake of what they strive to be). And that in turn means: beings that pursue those ends they take to be *good*. This point is crucial to understand so as to prevent the widespread misunderstanding that self-maintenance is simply preservation of one's bare life – "survival" or "mere existence". On the contrary. Because in the "is" of life the "ought" of fulfilment is always implicitly present – because life is essentially an ἔρως aiming at "the good", and because organisms stand in the self-relation (mediated by the world) of *caring* about their own *own being* – living beings (or agents) cannot *fully be* unless they *take themselves as such*.[114] That is, as *fulfilled* or *self-realized* in their actions. It is only because of this rift within, because of the possibility of failing to satisfy the *internal norm* in light of which the living being acts (and yet stay alive),[115] that one can also "become who one is",

---

112 Riccardi M., *Nietzsche's Philosophical Psychology*, p. 30.

113 GS 349.

114 The nature of that "taking" clearly varies among different forms of life (which also determines what they can count as fulfilling). Yet this basic purposive logic remains in place in all life. Nobody will be surprised to hear that there is a difference between a dog that takes itself as fulfilled and one that doesn't. That's why the dog plays, or tries to avoid pain. But even in plants, this purposive logic – although in a subtler, more subdued manner – remains present. After all, that's where the notion of *flourishing* (another fine equivalent of fulfilment) has its roots.

115 This is to say, the fundamental practically-normative distinction between "good" and "bad" (or even "evil", if we want), between what should and should not be, is purely formal and presupposed by the normative "is –

to use the Nietzschean (Pindarian) phrase.[116] And it is why Nietzsche rightly objects that: "Physiologists should think twice before putting down the instinct of self-preservation as the cardinal drive of an organic being."[117] For clearly, when an organism must, out of necessity, struggle to survive, rather than pursue activities expressive of its own nature, it is quite clearly reduced to a deficient existence, a state of distress.[118] In other word, the point is to ensure that proper conditions are in place so that the living being can, in its existing (which just is its pursuit of the good or fulfilment), not only exist in such pursuing but actually be fulfilled – by taking itself as such.

To sum up: Living beings are irreducibly purposive and normative agents. Beings that act for the sake of something (their subjectively perceived "good"); or, which is the same thing, in light of their internal norm that serves as the normative *measure* the organism uses to evaluate, navigate, and negotiate with the world. A world that is objective, existentially independent of the organism; a world that resists our desires (requires us to take means to our ends) and has the power to both confirm and falsify own subjective takings: square how things appear and how they actually are. Indeed only for such a being can the objective world be at the same *its* world; a world that is meaningful *for* it; a world the being *cares* about. For it is only *this world* in which the living being – as dependent, finite, and striving – can (and must) seek fulfilment.

---

ought" difference sketched above. For the movement of striving from the *is* to the *ought* (the good) can only make sense if there is something, at least potentially, "wrong" with the *is*. (And as we already saw above, that minimal potentiality of "evil" corresponds to the active prevention of it; or, which is the same thing, the active preservation of the good.)

116 See GS 335: "We, however, would strive to *become who we are* – beings who are new, unique, incomparable, who give laws to ourselves and create ourselves!"

117 BGE 13.

118 In GS 349 Nietzsche calls this the "incomprehensibly one-sided doctrine of the 'struggle for existence'" and astutely comments: "To seek self-preservation is the expression of an emergency, a restriction of the actual fundamental instinct of life."

When it comes to Nietzsche, we see that the will to power can be understood as simply coterminous with life. For life just is this drive to realize oneself in the world, in all possible shapes and sizes. And what's more, if, as we saw earlier, it is the task of psychology to make sense of life, Nietzsche's own peculiar definition of this discipline – "the *morphology* and *evolution of the will to power*"[119] – should no longer surprise us.

And yet, there is more to Nietzsche's notion of life. And it is precisely this part that, in the eyes of many, makes him a "dangerous" thinker. I speak of course of what Nietzsche takes to be the "moral" (or rather, *a*moral) consequences of his account of what it means to live – for Nietzsche does think there are some.[120] After all, the peculiar name of the "living will"[121], the "will to *power*", already betrays some of it. We haven't seen that side yet. In what follows, however, it cannot be kept out of view any longer. For it is only then that Nietzsche's philosophical project as a whole truly comes into its own. The reasons for that is simple: It is here that his philosophical *psychology* – what we have been engaged with thus far – finally meets *ethics*. As such, Nietzsche, the moral psychologist *par excellence*, can at last appear in his essence.

119 BGE 23 (emphasis mine).
120 Cf. GS 355.
121 GS 349.

Part Two

V. The Amorality of Life

As we saw in the last chapter, the notion of the self-affirming ("power-seeking") drive is the centrepiece of Nietzsche's account of life. The naturalistic replacement of the Platonic-Christian *erós*: the primal striving towards the universal *summum bonnum*, the one true end of life. Of course, Nietzsche turns this logic upside down. Rather than conjuring up some independent, objective Good or Truth towards which we must aspire, he makes do with nothing but our desires and drives (our physio-psychological constitution and environment) whose ends – simply by virtue of being those ends – we in turn call "good". For Nietzsche might well ask: What could an alternative naturalistic theory look like? How could it explain feelings, thoughts, and action? How could an organism perform anything without some *internal* preference for the outcome? It couldn't. The said action would simply not occur. In that case, however, the agent can hardly be some neutral bystander in this process. On the contrary, it must be an interested party – and that is the same as: self-interested party.

It is precisely on this "internalist" logic that Nietzsche's "psychological" critique of morality is built. For Nietzsche believes that just with this account in hand we can already show morality – at least as traditionally conceived – as a constitutive impossibility. Of course, first we should probably inquire what the real target is. After all, "traditional morality" can be a promiscuous term. Fortunately, the answer here is surprisingly simple. Nietzsche reacts to the common wisdom (accepted, among others, by Kant and Schopenhauer) that says a moral or virtuous act is one that is performed selflessly – out of *selfless motives.* And for Nietzsche, this is enough to get

the demolition going. For he is convinced that his internalist account of motivation, which he takes as the only one naturalistically viable, simply leaves no room for (real) selflessness or altruism whatsoever. This is most evident in *Human All Too Human*, where we find Nietzsche's strategy at its most straightforward.[122] His palpable wish to naturalize agency – that is, the denial of all supernatural motives and his will-to-power "drive psychology" – even leads him to utter the sweeping claim that, insofar as moral action means action done solely for the sake of another, *there has never been a single moral action*.[123]

> "Never has a man done anything that was solely for others and without any personal motive; indeed how *could* he do anything that had no relation to himself, that is, with no inner compulsion (which would have to spring from a personal need)? How could the 'I' act without the 'I'?"[124]

As we can see, Nietzsche is not just making the trite psychological observation that people tend to act in an egoistic manner. Invoking logical necessity, he tells us that "*the whole concept* of 'selfless action' evaporates upon closer examination."[125] That is, there could never possibly be anything truly altruistic. So, *that* is what's at stake here – and what we must hence examine.

Still, the first point to make here should be an obvious one: "Selflessness" cannot mean an *absence of a self*. How could *I* act without the "I" indeed! This might sound obvious, but it is important to keep in mind so as to avoid falling into some predictable traps further down the line. And it is relevant for Nietzsche's deeper point. A point he begins to make when when he

122 Later on, Nietzsche complicates things, as we shall see shortly; but the basic line of attack (i.e. challenging altruistic motives) remains unchanged.

123 See mainly D 148.

124 HH I.133.

125 *Ibid*. (emphasis mine).

confronts the reader with some paradigmatic "selfless" acts (such as a mother's self-sacrifice for her child, or an author's undiluted concern for the truth, rather than, say, personal fame or wealth) – and asks: Are these *truly* selfless?

> "Is it not rather that in all these cases, one is loving *something of himself*, a thought, a longing, an offspring, more than *something else of himself*; that he is thus *dividing up* his being and sacrificing one part for the other?"[126]

Nietzsche's meaning here seems simple enough: We cannot get "outside ourselves". Thus, when we look closely, at the deepest level the very dichotomy between the selfish and the selfless simply dissolves. The question, however, is what to make of this. For should we say that this "dissolution" – due to the fact that in every act the self is inextricably implicated – allows us to see that altruism, love, charity, philanthropy, and indeed morality as such are in fact merely another species, or "sublimated" expressions, of some original egoism?[127] That while morality may claim selflessness and love as its proper values, the truth of those values lies not in their "essence" but rather in their "usefulness"?[128] And that, consequently, rather than about altruism, morality is usually really about its opposite?[129] In other words, what exactly does Nietzsche think this "deconstruction" of his accomplishes? What significant consequences can we draw here, if any? For one, it seems obvious that the notion of egoism becomes intelligible and significant

---

126 HH I.57.

127 See HH I.1 where Nietzsche says that "strictly speaking, there is neither unegoistic action nor completely disinterested contemplation, they are both just sublimations in which the main element appears almost dissolved and proves to be present only to the finest observation."

128 HH I.133.

129 Spinoza's remark, that whoever loves God truly cannot expect to be loved by Him in return, comes to mind (see Proposition 19 in Book V of the *Ethics*), for it applies equally well in the purely human sphere – a fact that wasn't lost on Nietzsche (given that he never tired of pricking the "Christian" conscience).

only when counterposed to its opposite. One cannot hope to simply erase one side and keep the other. And pointing out the necessity of the Self's involvement also seems like somewhat of a misdirection as well.

Now, in a way, that is exactly Nietzsche's point. He does want to show us that this "selfishness" is something unintentional, unavoidable, primitive; something *constitutive* – and thus perfectly innocent. That the opposition between the two poles is far from absolute and "given", and that it is ultimately us who decide, presumably on pragmatic grounds, where the line is drawn.

> "Egoism and its problem! The Christian gloominess in La Rochefoucauld, who drew it out from everything and thus thought the value of things and virtues *diminished*! *Against him,* I first sought to prove that there *could be nothing else but egoism* (...) that love is an expression of egoism, etc."[130]

And yet, as we can see, Nietzsche does want to have it both ways in the end. Or rather, *only one way*, in that he still wants to prioritize one side *over* the other; namely, give egoism an ontological primacy and make it the "main" (and really, the *only*) element.[131] And this means we cannot avoid posing our question again: How could any morally significant "egoism" possibly follow from something as obvious as the denial of *self-less* action?[132] Can we truly base a critique of (any) morality on something as basic as this? If so, what would it need to look like?

---

130 WP 362 (emphasis mine).

131 See HH I.1, where Nietzsche openly declares (and assumes) that egoism "came first" and everything else is merely a "sublimation" that preserves its basic logic.

132 That is, if Nietzsche's point is to have any moral purport at all, he must provide a demonstration of how this constitutive self-relation, *as such*, makes "moral" (or "altruistic") acts impossible.

To start off, it seems plainly obvious that the only moral theories that could ever suffer from this criticism – could be ruled as *a priori* inadmissible – would have to be deeply incoherent themselves. That is, would have to assume the impossibility of *self-less* action. But what theories do? There may be some spiritual teachings that see the Self as enemy, but even there, the point is rarely pressed to the extreme. Nietzsche's point is surely subtler that this. And as we shall see later on, it is. For what he means to say is this: That while traditional morality might not avow such constitutive impossibility as self-less action explicitly, it nevertheless stands on wrong assumptions about *what can and cannot motivate us*. And because of these false premises, it also ends up – in a cruel and moralistic fashion[133] – prescribing the impossible: imperatives we are simply unable to follow; "oughts" that do not imply "can"; and thus *actions no Self could perform*. But let us not get ahead of ourselves. Thus far we have seen only that if we are to act, we must be internally motivated to do so. And that this translates to: the reason why we *do* something is because we take it to be "good" (or "desirable") – which just *is* the reason, or motive, for doing it.

On its own, however, this tells us nothing. Or rather, close to nothing; for by denying the possibility of *self-less* action, Nietzsche has indeed brought into focus the logical condition for the possibility of any action whatsoever: the purely formal fact that all my interests, whatever they may be, must in fact be *mine*; that the motive for my acting must be *my* motive; that for an action to count as a genuine *action* – that is, a minded "agential" expression of my-self (something *I do*), rather than a mere event (something that just *happens to me*) – it must issue

133 As Nietzsche notes (albeit in a slightly different context in) GM II.6, "the categorical imperative reeks of cruelty". And of course Kant, with his "noumenal" realm and the pure practical reason and its postulates is (along with Schopenhauer) one of Nietzsche main target on this score. GS 335 expresses a similar thought.

from and aim at this very Self. In other words, what we see here is that for me to care about anything, there must always exist an internal connection between *that what I care about* on one hand and *myself* on the other. (Hardly a great metaphysical breakthrough.) This way, far from demonstrating some morally significant "psychological egoism", we only come to see in slightly different terms what we've already seen above: That the very activity of living, the self-relation of striving towards the good, is what *grounds the possibility of things mattering to such beings in the first place*. And that things can therefore only matter to a being that cares about its own being; a being to whom it matters that it be the sort of thing it is. But this merely shifts the decisive question. For what we need to know is precisely what this "it" is. And so we must now ask the question: What is this Self? What makes me *me*? The self that *I* am. – It is here that Nietzsche's deconstruction of said dichotomy will have to be decided.

The problem is, it appears to crumble straight away. For when Nietzsche observed that in every action one affirms merely *something of oneself*, his point was precisely that *everything* in the world is only accessible to us as, in a sense, included "within us". For in none of your actions can you ever truly step "outside" yourself, or "beyond" yourself (as it wouldn't be your action then). Instead, in everything you care about, in every choice you make, the end you aim at is always something with which you *identify* – that is, something in which you "see yourself"; something that you recognize as *good for you*; and thus a "part of yourself", a part of who you are, a part of your being. Indeed absolutely everything "out there" – if it is to be anything to us – can never be "other" as something absolutely alien (in which case it would be nothing to us), but merely as something we relate to – and through which we relate to ourselves. In fact, through which we

constitute ourselves. Thus, what Nietzsche really shows us here is that we all acquires our specific identities – and thus come to be concrete Selves – only by being included in the world.

For what else is there? Who is it that remains once you strip away all of the associations and affinities that shape you; that make you *you*? The abstract, universal (that is, *impersonal*), "transcendental Ego"? Nietzsche – someone who never tires of challenging the superstition of the neutral "subject behind the deed",[134] as well as of emphasising the irreducibly relational, processual nature of reality[135] – knows all-too well that finite beings are not some self-enclosed atomistic substances, but rather dynamic "acts" of relation to what is other than themselves. With us as humans it seems especially obvious that we simply cannot be persons *qua* some pure "individuals". Our social being and social identities are not "add-ons", optional extras sitting *on top* of our "individual being". Instead, *individuality itself* is only possible dialectically – as internally related to, or necessarily dependent on, the external world of other beings, both inanimate and animate.

The only pertinent question should thus be: *Which of these* "parts of ourselves" (that are nevertheless other than ourselves) ought we to adopt and prioritize? To what practical projects or existential ends should we – both individually and collectively – commit? Only here, on this level, does the distinction between the selfish and selfless, as well as moral and immoral, become intelligible – become *a question*. For only here can we start answering which of our commitments ought to count as worthwhile. But why on earth should that in any way degrade them?

---

134 See GM I.13.

135 See for example WP 557–9.

Nietzsche seems to think that, because we are always so self-related, it is impossible to treat these "othered" parts of us as truly, *inherently*, valuable. For he holds that, if in every action each organism always strives toward its own good, and in a way thus "aims at itself", then this good will always be something that, by its very nature, *excludes* the possibility of such "altruistic" action. That the self-affirming will or "drive" of living is necessarily incompatible with caring for our specific commitments unconditionally – *whatever they may be*. And that, of course, includes caring for others; valuing and treating them as *ends in themselves.* Yet why should this follow? This runs into the very same difficulties we've just been through. Our consideration of the socially, normatively constituted Self dissolves this cynical scepticism entirely. That is, if Nietzsche wants to say that others can only be *mere means*, then we must ask: *means to what, exactly*?

Here Nietzsche's own gloominess seems to bring us dangerously close to where our paradigmatic model of explaining action becomes that of *external*, rather than internal, purposiveness.[136] However, on this model, just like on the *mechanistic* one, infinite regress ensues at once. One finite end is "explained" by yet another finite end; and as a result – *nothing* is explained. All explanation is simply deferred. Which is to say, it is not really (or not just) the regressiveness as such that is the problem here. The real issue is that, because of this regressiveness, we cannot even understand *any* of the individual "steps" – the very idea of caring about something, of having a purpose and desiring that some state of affairs be brought about –

136 This is essentially the Humean or empiricist theory of action and subjectivity, according to which conduct is ultimately determined by subjective, psychological *desires:* intrinsically motivating non-cognitive, non-conceptual, causally elicited stimuli that "move" the organism to bring about certain state of affairs. On this account, reason thus takes on the role of an instrument; the servant of the passions, as Hume famously tells us. As such, reason is "employed" *by* the desires for the task of selecting the appropriate leading to the desired, causally pre-determined end.

in the first place. And the problems don't stop there. For even if we did posit some "final desire" to stop the regress,[137] this model still couldn't explain the very relation between ends and means; or the very activity of discriminating between means and obstacles. Because it treats desires as effects of external causal stimuli (which "set us" specific goals) – that is, as something *independent* of normative, rational evaluation – it remains a mystery how we could ever come to understand some means *as appropriate* for bringing the desired state about. Put differently, how can a causal impingement ever come to *specify* what it would mean to achieve the desired aim or object if it is completely independent of all conceptual content (as that which is required for all specification)?

External purposiveness thus fails not only as an account of subjectivity (as pure instrumentality) but objectivity as well. Because it cannot provide the norm that could distinguish between "good" and "bad", appropriate and inappropriate, or rational and irrational ways of relating ourselves to the object, and thus, it cannot account for the very determinacy, the form, of the object itself (what "matter" would count as fitting or appropriate for it).[138] As such, all such merely "subjective", external purposes must therefore always be "enveloped" by internal ones – the reasons one has for valuing something on its own terms. That is, while you may care for someone else because you desire, say, the pleasant company the other person provides, that desire must itself be grasped as the reasons one could have for caring for the other as a true

137 As Hume notes we must. "It is impossible there can be a progress *in infinitum*; and that one thing can always be a reason, why another is desired. Something must be desirable on its own account, and because of its immediate accord or agreement with human sentiment and affection." (Hume D., *An Enquiry Concerning the Principles of Morals*, Hackett Publishing Company, 1983, p. 87.)

138 Again, like mechanism, external purposiveness treats function as accidental and "merely subjective": as if we impressed on houses (as if indifferent matter) our contingent desires to live in them. But there is no house to be understood without its purpose. And what's more: there would be nobody to understand them if internal purposes (such as dwelling) weren't constitutive of us as living agents capable of grasping them.

purpose, something valued for its own sake – a "truly infinite" end that must be sustained to be what it is. Not merely an instrument, but also the end of solicitude as something that is appropriate to for the kind of being that we are. On this account, the activity of caring would thus be a *constitutive* part of what human beings *ought to* do.

In any case, just like with the isolated drives of Chapter III, we can make sense of behaviour only when we understand our desires, projects, and purposes as manifestations of the transcendental, existential unity of striving toward the good – as attempts at attaining this "ought" of fulfilment: the final end of all purposive life and its formal measure of what counts as worthwhile. But of course we already know that Nietzsche does have a concept that seems to play just such a role. According to him, this "highest good" is *Power*. Thus, if Nietzsche still wants to affirm that such an end should be incompatible with altruistic care, we must now ask how – exactly – we are to conceive of it. This means we must re-examine how this "power" truly functions. That is, what explanatory role it has, no matter its concrete content, but merely *qua* such an end. In particular, the question that concerns me most is this one: Is power a "substantive" end, something out there in the world to which we might point (one "thing" among others or something akin to the "aim" of eating/nourishment)? *Or* is it the *formal* end constitutive of the structure and the dynamic of living as such – the principle that makes it possible for us to understand action, including our distinguishing between ends and means, in the first place? If it is a formal end, why should it *a priori* exclude anything? And if it is a mere substantive end, why should it be absolute? The whole thing seems extremely implausible.

Of course, just about everywhere Nietzsche suggests that power should be understood formally.[139] After all, what is the whole point of the "will to power doctrine" if not to show us how it can explain, without introducing any mysterious "skyhooks", not merely the putatively normal behaviour that doesn't puzzle anyone, but also the seemingly senseless, self-denying acts – behaviours such as asceticism, masochism, martyrdom, suicide, self-sacrifice, or indeed, compassion, pity, and altruism – that we might (mistakenly) think do not aim at fulfilment or "well-being" at all?[140] Such power can hardly be *a* good; something we take to be fulfilling because of something else. On the contrary, it is clearly supposed to be precisely *that which specifies* what our particular should be.

Some commentators suggest we read the will to power as a second-order phenomenon;[141] a by-product that necessarily accompanies all of our desires – themselves oriented at specific "first-order" ends *distinct from power* (say, nourishment or sex as "relative" ends or "aims"; or more specifically, this or that particular object which holds the promise of their attainment). This, I believe, gets things *almost* right. However, it can also distort them rather heavily. First, it must

---

139 In some places, Nietzsche does oscillate between power understood as (1) the essence or true nature of life as such – something which, in reality, the organism at all times inevitably seeks (which yields the formal reading); and (2) merely one drive (*an* "aim") among many, albeit perhaps the most powerful one (which would yield the substantive reading). The few passages suggesting the latter (such as GM III.18) are swamped by the former (see footnote 106); yet such an inconsistency should worry us greatly, as it suggests that Nietzsche himself was confused about this crucial point. In any event, the point is this: Even if "power" should be one end among others, it would still need to be situated within the formal, internally purposive structure of life that necessarily entails the final end of the organism's "good" – and that obviously *cannot* be substantive.

140 Nietzsche wants to show us that we don't need to (and shouldn't) posit any miraculous principles of purely altruistic pity, empathy, nor some unknowable "noumenal" realm from whence all truly moral (or for that matter, immoral) action springs. On the most general level, *all* of our actions are perfectly explicable in exactly the same terms – as expressions of the one universal self-affirming will (our logic of striving toward fulfiment). The only reason why certain behaviours seem senseless to us is because we are not subject to the same conditions as their "subject". But whatever the content, the basic internalist logic of action must remain the same; otherwise the deed simply wouldn't take place.

141 For example, see Richardson J., *Nietzsche's System*, p. 21–3.

again face the problem of pulling apart life's purposive unity. And second, once we understand the will to power formally, such that "power" just *is* our word for fulfilment, then on this account power just *is* what distinguishes the desirable from non-desirable, the satisfying from the non-satisfying, the good from the bad. For in everything satisfactory, one inextricably and unavoidably attains some measure of fulfilment (here power) – as *that is just what satisfactory means*. Given this, however, to desire anything *is to desire power*.[142] Thus, while such "power" can be understood as a by-product of sorts (in the sense that, again, it isn't an empirical thing with observable properties "out there" but a "final perfection" (τέλος) that *emerges* for us when our justification of a purpose and the purpose itself become one),[143] it is no mere secondary result of fortuitous coincidence, *but the constitutive end of any action whatsoever*. And as such it must be afforded explanatory primacy.

Substantive ends simply do not function like this.[144] If power were just one end among others, it would be a contingent end, an optional end. One for which the organism may or may not strive – *depending on whether it will matter to it*. Which is to say, whether the organism will take it as valuable, as meeting the internal norm or standard of fulfilment (this being the *formal*

142 There is another distorting way to understand the "power as a by-product" thesis. According to this view, what one always unconscionably wills is first and foremost *the struggle* itself – the overcoming of the gap between the (current) state of *lack* and the (desired) state of *fulfilment*. But of course it is an analytic banality to say that to truly will an end one must also will the means; and that these means can thus be understood as an ability or "power". The world certainly didn't need to wait for Nietzsche to point this out. Also, how strange to suggest that one should desire to be in the painful process of struggling with resistance. Nietzsche thinks one desires power – to be overcoming – but for that to be the case, for the struggle to show up as overcoming, it must clearly be itself fulfilling (in a sense, we must *have already overcome*; the struggle must have lost its negative status).

143 That is, when the activity is an end in itself; something done for its own sake; something inherently fulfilling.

144 It has been noted that by wanting to explain everything the will to power explains nothing. But of course that is precisely what one expects from a formal, constitutive principle. For "rules" like these don't explain why someone chooses this over that; they make it possible to understand the phenomenon of choice (*as choice*) in the first place (and thus to distinguish it from, say, purely mechanical differential responsiveness).

principle in light of which all contingent purposes or commitments become intelligible as such purposes in the first place – hence the explanatory priority). And with this, let us return back to the main topic of altruism and our reasons for action and summarize what we've found.

Why do we pursue ends? We pursue them *because we care about them*; because they matter to us. A mother who genuinely cares about being a mother doesn't care because it brings her *something else*; she cares because she loves her child. Being a mother is, to her, an end in itself. But then – and this is the crucial part – so is the child. Obviously, the two cannot be separated, given that caring for her child just is what being a mother means. In that case, however, it would be patently absurd to claim that the mother doesn't "really" care about the child. That the child is necessarily *but a means* or "instrument" to her "personal interests": whether it be power or some other such distinct purportedly "egoistic" end.[145] No. Fulfilment, as conceived of here, is a formal principle; it is the measure we use when we evaluate whether this or that "content" ought to count as inherently worthwhile, as an end in itself, and thus not something *other than* the content it "enmatters".[146] Therefore, once we grasp "power" formally, as a true final end, *then there simply is no problem* (except, perhaps, the peculiar word choice). The entire issue of the possibility of altruism vanishes. For while fulfilment certainly cannot be separated from the self (as self-less actions are no actions), it also cannot be something entirely separate and self-standing merely on "our side". On the contrary, true fulfilment resides precisely *in* – or *coincides with* –

145 In fact, if we isolate the particular relation of love, we may say that what makes it (true) love is precisely that the self in question is necessarily and irreducibly experienced as *us* (such that the being of the other is a *constitutive* part of myself). If Nietzsche wants to deny that this in fact possible – that in principle, said mother cannot truly care about *anything but* power (or pleasure or some such end) – then he merely traps himself in the regressive model worked out above.

146 Because a final end is not some physical entity in the world (we cannot point to it or observe its properties) but a form of some thing, it is thus essentially "enmattered." Which is to say, once again, that it is *nothing on its own.*

those activities that are worth doing for *their own sake*; activities, that is, where means and ends merge with one another so completely that even the most difficult parts of the activity in question show up as fulfilling and inherently motivating. Which is why even the sacrifice of my life for someone else's can show up to me *as good* – as Nietzsche well knew.

Of course, this mustn't be misunderstood to mean that there are guarantees. What matters to us is never truly settled. Caring about being anything (a parent, a son, a spouse, a teacher, an author, a friend) is, as we have seen, a commitment that must be sustained, and thus something that can break down (even deteriorate toward being a mere instrument to something else, no doubt). Likewise, none of this is to dispute that all around a myriad of *ingenuine* cases of "care" continues to exist. People can pretend to love someone even thought they really care about something that doesn't involve the other person – as *a person* – at all. In such cases, the other's being – their fulfilment – is not a constitutive part of one's own, but at most an accidental by-product. And yet, *none of this* has anything to do with the self-relating character of living agency. It has everything to do with hypocrisy, with solicitude being only a means to something else (and thus not being real solicitude),[147] but also with self-deception, our own fallibility, and the sheer reality of the economy of finite life that forces us to choose, focus our attention, and prioritize one commitment over another (create a "rank order" of values, we might say). But that doesn't mean that acting on behalf of someone else *has to be* some alien purpose, or a mere means toward some external "selfish" end. It *can* be an expression of a commitment in which we truly recognize

147 For example, when being a spouse is really about prestige (say, in the case of a husband who has found himself a trophy wife); or when the good and caring capitalist turns out to care about financial gain after all; or when the philosopher committed to truth sets out to philosophize not with truth in sight, but rather lasting fame; and so on. On the other hand, notice that there is nothing about genuine care that necessitates excluding these things. Prestige, profit, or lasting fame may very well follow – although accidentally – follow.

ourselves; a purpose that makes up an essential part of our self-understanding – and thus ourselves, period. As such, all that can possibly matter is simply whether our purposes are or aren't ends in themselves.[148]

But if that is so, why should that exclude altruistic care for others? Isn't the commitment to someone else's flourishing just one possible end among others? One that we may or may not take up? At most we can say that it is always an open question whether specific persons actually have good reasons to care for (certain) others; and that the normative, ethical questions of whether they should, in what capacity, etc. are not ones that can be easily answered *in abstracto*, without considering the broader context (such as the particular social conditions under which we live). But there is absolutely nothing that should make altruism a constitutive impossibility. Nietzsche's attempt "to prove that there could be nothing else but egoism"[149] has thus failed. If it is to continue, an altogether different strategy will be needed.

148 Commonly, selfishness and selflessness are taken to be dependent on how the "doer" figures in the motivational structure of the given action (for instance, me doing something for my own financial gain, or a family member, is taken as less altruistic than, say, donating money to children in Africa I have never met and am not "personally" invested in.) The point of the preceding analysis was to show that this is misleading (if not completely orthogonal to the real issue). For I am always implicated in whatever it is I do; and cannot ever truly deny or disown myself. And as a human being, it is only through relating to others that I come to be my self. The only thing that *can possibly matter* (not pragmatically of course; merely with respect to the distinction we are investigating here) is thus simply whether I *care* – truly care – such that the other involved in my action is taken and treated as an end in itself – a *constitutive* part of me. As such, the key measure of selflessness turns not on whether *I* am involved (I always am), but on *how the other is involved*.

149 WP 362.

VI. Anti-Rationalism

In the last chapter, I attempted to show that altruism – non-instrumental treatment of others – should be a perfectly real possibility. However, that account was thoroughly depended on *normative, rational self-identification* (that is, on "conscious ascriptions"). On our ability to intentionally, or freely,[150] *set our own ends*; and on the "space of reasons" having a say in determining our course of action. And as will become clear bellow, Nietzsche's final strategy will consist of challenging just that. For it will be shown that he believes our conscious ascription, including (potentially altruistic) intentions, to be – in a way – *inert by-products*. As such, we will finally arrive at what I alluded to in the previous chapter: The fact that Nietzsche's whole "internalist" account and polemic ultimately targets the (in his eye's mistaken) notion that intentional, rational reflection or conscious deliberation has the power to "move us" – to determine and guide our conduct. In other words, like Hume before him, Nietzsche will want to deny *practical reason.* As a results, the efficacy of the rational will, and in a way, self-consciousness as such – the very fact *that we are self-conscious agents* – will have to be reevaluated.

In Hume's texts reason is, to be sure, denied the ability to set our ends, which belongs to the passions. Yet it is assigned the subordinate role of the obedient servant, cunningly figuring out how best please its master.[151] As it happens, I already considered this conception, under the

150 For it is intentionality – the fact that an action embodies some (at least minimally conscious) judgment – that allows us to distinguish between actual actions or deeds (something I do, such as my raising my arm) from mere events (something that happens to me, such as a twitch in virtue of which my arm gets raised). Unless we preserve intentionality, as something intimately linked with both thought and freedom, we deprive ourselves of the means to account for a whole range of phenomena; chief among them: life and action (and everything they entail – which, however, is arguably everything, given that the possibility of objective knowledge constitutively depends on such living, purposive agents).

151 "It appears evident, that the ultimate ends of human actions can never, in any case, be accounted for by reason, but recommend themselves entirely to the sentiments and affections of mankind, without any dependance on

rubric of "empiricist theory of action", in the context of my critique of "subjective", external purposiveness and instrumental explanation in the last chapter. And there I came to the conclusion (although I haven't expressed it in so many words) that such theories must ultimately affirm something like what John McDowell has called "the unboundedness of the conceptual";[152] as well as that self-conscious reason *can in fact set its own ends* (because that's just what it is) – that who we *take ourselves to be*, our self-conception, is indeed an essential part of *who we in fact are*. But technically, there is another option: We can try to get rid off reason altogether and thus end up with some kind of epiphenomenalism or eliminativism.

In the rest of this chapter, I argue that even despite Nietzsche's apparent conceptualism, he is nevertheless committed to certain "overriding" theses that give us no choice but to place him squarely into the latter camp. The camp that believes our lives fundamentally governed – that is, determined – by forces that are by themselves non-rational, that is, by *unreason*; and that the efficacy of rationality is thus merely a subjective "user illusion", a fictitious by-product of our internal, physiological make-up and our peculiar perspective on the world. The result of Nietzsche's going down this route, I will show, is a picture that does away with just about everything significant we've been through so far: purposiveness, self-determination, life, and agency as such. Critiquing this move itself will be the main purpose of my last chapter. But first things first. To fully appreciate what the stakes are, we must see what it is, exactly, that Nietzsche wants say, and what he wants to deny. Let us therefore take a step back, return to the account of the last chapter, and see where Nietzsche departs from it; and where we must thus locate the

---

the intellectual faculties." (Hume D., *An Enquiry Into the Principles of Morals*, p. 87.) And of course: "Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them." (Hume D., *A Treatise of Human Nature,* Clarendon Press, 2007, p. 226.)

152 See mainly the second lecture/chapter in McDowell J., *Mind and World*, Harvard University Press, 1996.

ultimate breaking point. In the process, I will also briefly sketch what the Nietzschean alternative looks like.

~~~

In the last chapter we examined Nietzsche's attempt to show that traditional virtue and morality *qua* selfless action should be discarded (as sanctimonious, moralistic egoism of the resentful sick and weak, posing – in nothing but a hypocritical power-play of its own – as "altruism") in favour of the "healthy" egoism of the noble master;[153] all of which was supposed to result from the unreality of truly selfless behaviour – acts performed solely for the sake of another. However, in thinking through and developing and account of what it means to be a Self, we came to conclude that Nietzsche's primary strategy fell short of the mark. The question before us now is therefore this one: Provided that Nietzsche still wants to deny altruisms significance and draw his "amoral" conclusions, what part of this account does he reject?[154]

There are good reasons to suspect that by the time of writing *Daybreak*, Nietzsche came to realize that attacking the possibility of altruistic motives won't do. For in that book, the proper beginning of his sustained critique of morality, his strategy of going about this pervading problem shifts: Whereas in *Human, All Too Human* he tried to tackle the problem of altruism's very *existence*, now he starts scrutinizing its *value*. But of course once we start evaluating what

153 See BGE 265.

154 There are some who believe that Nietzsche doesn't actually reject anything about it. For instance, Bernard Reginster has defended a reading that Nietzsche, far from proving but even advancing psychological egoism, comes to realize our intentions can in fact be altruistic – precisely because they can be non-instrumental toward others. Of course as we've seen this is all quite true. Yet as we shall see, as an *interpretation* of Nietzsche it misses something crucial. (See Reginster B., "Nietzsche on Selflessness and the Value of Altruism," in *History of Philosophy Quarterly*, 17(2), April 2000, pp. 177–200.)
~~~

altruistic acts are worth – how desirable they are, who they benefit, in what way, etc. – we accept that they in fact exist.

At this point, much of Nietzsche's thought on the matter turns toward a pragmatic and utilitarian re-evaluation of the actions hitherto "decried as egoistic" – to restoring them from their "immoral" status and curing men from the guilt and bad conscience associated with performing them.[155] This way, as Nietzsche says, we shall rescue the majority of life and action from "its *evil appearance*"; for when one "no longer regards himself as evil, he ceases to be so!"[156] This dimension, however, doesn't really interest me here. Yet, putting Nietzsche's prescriptions aside, but holding onto his critique of morality as such, there is another, philosophically much more significant aspect that underlies, even *grounds*, the whole enterprise. An aspect inextricably linked with Nietzsche's sceptical thoughts regarding reason or consciousness (some of which we had already seen in chapters I and II). To get us on the right track, let us begin by considering just in what sense can the altruistic act – whose existence Nietzsche now accepts – be said to "exist"? In what way is it supposed to be "real"? The following passage contains the answer.

> '*No longer to think of oneself*.' Let us seriously consider: why do we leap after someone who has fallen into the water in front of us, even though we feel no kind of affection for him? (...) The truth is: in the feeling of pity – I mean in that which is usually and misleadingly called pity – we are, to be sure, not consciously thinking of ourselves but are doing so *very strongly unconsciously*.[157]

155 D 148: "Our counter-reckoning is that we shall give back to people their good will towards the actions now decried as egoistic and so restore their *value – we shall rob them of their bad conscience!*"

156 *Ibid*. (my emphasis).

157 D 133

What Nietzsche tells us here effectively amounts to this: Selfless acts do occur, yes, but only in the sense that people ascribe to themselves – *consciously* attribute to themselves – such (altruistic) *motives*. When we pity someone, we are not *consciously thinking* of ourselves. – And there's the rub of course. For what we come to see is that Nietzsche actually thinks these conscious ascriptions, motives or intentions, as such, *fundamentally erroneous*.[158] That is to say, while we may think that we know why we're doing what we're doing in, say, jumping into water to save a complete stranger, what we *think* we know and what we *really* know are, for Nietzsche, two completely different things. Even if, from our conscious, practically engaged, first-person perspective, they always *seem* the same. This is why, when few sections later he once again returns to the puzzle of altruistic action and asks, "And what is it then that is so *named* and that in any case exists and wants explaining?" – he concludes: "It is the effects of certain *intellectual mistakes*."[159] Thus, while we do get an affirmation of altruism's existence, that altruism is immediately pronounced *inert* – inert because it only exists *as an illusion*. In other words, aware as he is that on our normative model of the self one cannot really question selflessness and the foundations of traditional morality, Nietzsche *rejects the entire model*. What he gives with one hand, he takes away with the other.[160]

---

158 With this acknowledgment, Nietzsche's campaign against morality may now properly begin. For now he can avail himself of the claim that people can genuinely *believe* they act for moral reasons (just as they can genuinely *believe* in the loving God) – instead of having to awkwardly say that really they believe something else – but that their beliefs suffer from a kind of "presupposition failure". That is, the assumptions on which these actions stand are themselves untrue. See mainly D 103: "I deny morality as I deny alchemy, that is, I deny their premises: but I do not deny that there have been alchemists who believed in these premises and acted in accordance with them."

159 D 148 (emphasis mine).

160 Of course, in a way, he doesn't really have a choice. For how could we deny altruism *as such*? After all, the pangs of conscience – the very fact that sometimes we do seek selfless "purity" in our motives (however impure it in fact might be) – are a testament to its (phenomenal) reality. As an experience of one's own intentions or motives – altruism simply cannot be falsified. At most, it can be deemed epiphenomenal – which is what

Accordingly, when Reginster defends Nietzsche as not really advancing egoism after all,[161] he misses that Nietzsche's war on morality is waged on two fronts: That apart from (unsuccessfully) challenging the purity of our perceived or conscious motives Nietzsche also wants to claim that, when it comes to the question of true motives for action (irrespective of whether we experience them as selfless or not), *conscious motives do not play a decisive role at all.* But surely that is of the highest import here. For it is precisely these intentions on which the whole theory of *self-conscious identification* stands.

Then again, it is quite understandable that Reginster wants stop where he does. After all, Nietzsche often does speak in a way that suggest or assumes both the importance of motives and our ability to discern them. Otherwise, what would be the point of all his sustained efforts to unmask the real motivations for why certain people act and believe what they do?[162] Yet the simple fact of the matter is that such an interpretative move comes at the price of ignoring some of the most philosophically significant things the "scientific" Nietzsche has to say – as will become clear bellow. Therefore, thought I agree with the thesis on philosophical grounds,[163] I

Nietzsche does.

161 Reginster B., "Nietzsche on Selflessness and the Value of Altruism," in *History of Philosophy Quarterly*, 17(2), April 2000, pp. 177–200.

162 As Reginster himself notes: "Nietzsche takes for granted the view that the altruistic value of an action depends essentially on its motives. This comes through clearly, I think, in the very nature of his criticisms of apparent cases of altruism: they almost always consist in exposing motivational deficiencies, by showing how they are ultimately animated by covered selfish motives." (*Ibid*., p. 178.)

163 Although Reginster still thinks it makes sense to speak of "selfish altruism", as in: "Nietzsche's paradigmatic mother is selfish in the sense in which her ultimate concern is with her own happiness, but her care for her child possesses altruistic value nonetheless as the value she places on her child's happiness does not derive from its relation to her own. (...) Accordingly (...) certain selfish actions do possess altruistic value, and Nietzsche is right in claiming that there is no (necessary) opposition in value between egoism and altruism." (*Ibid*., pp. 196–7) – I on the other hand have tried to show that this is a vacuity. Even Socrates or Aquinas explicitly affirm we all strive towards our good. Should we say *they* believe we must be "selfish"? There would be no striving, no action, no life, and no possibility of selflessness if this weren't the case.

want to pursue Nietzsche's critique further to evaluate whether it has the potential to do what Nietzsche (*contra* Reginster) thinks. And that means turning to those very "intellectual mistakes" concerning our misconceptions about reason and what sorts of beings we are. For only this way, Nietzsche holds, can we uncover the true blunder, the true cause of the epochal confusion we call morality.

~~~

At this point, we should have a pretty good picture not only of what it is that Nietzsche wants to accomplish, but even of how he means to do it: Our fundamental question is how to best explain organisms; that is, how to best account for their seemingly irreducibly purposive nature (including our own reflective "rational though"), their ability to act in such a way that seems to require different explanatory strategies than mechanical interactions of inorganic bodies (*qua* truly inert matter).[164] This is a problem that had puzzled Nietzsche from the beginning. His doctoral dissertation was originally supposed to be on the concept of the organic since Kant,[165] and in his early unpublished text "On the Origins of Language" Nietzsche writes that the real problem of philosophy is "the unending purposiveness of organisms *and the unconsciousness in*

164 The organic, purposeful, desiring being – in contrast to the inorganic one – is such that it has a "good". Accordingly, organisms cannot be understood on the mechanistic model for there are only certain ways their desires may be satisfied, and those depend on what the (subjectively) desired object (objectively) turn out to be. In other words, there is a way to fail and succeed for an animal. Light, on the other hand, cannot fail when it refracts; and rocks don't care whether someone breaks them because there is no way they *should be* (other than what we subjectively project onto them). Or to put it in a more general way, valenced, intentional awareness and responsiveness cannot be collapsed into such mechanical "responsiveness" (and efficient causality only) because then intentionality would be completely emptied out of content. And that would mean that we'd have no means of distinguishing between actions and events.

165 See Kant I., *Critique of Judgment*, §§64–66, where Kant presents his conception of the organism as irreducibly (internally) purposive – no matter the organism's intellectual capacities. His point, which is also Nietzsche's point, is that there are certain kinds of object that *we* simply cannot understand without appealing to the notion of purpose.
~~~

their coming to be."[166] I highlight the second part because it makes clear that Nietzsche's foremost issue was not merely with the notion of purpose as such, but perhaps even more importantly with purpose as something the living organism doesn't necessarily represent "for itself", such that it would have to be self-aware of its reasons for behaving the way it does.

The motivation for wanting an explanation of such phenomena is quite understandable. Again, we might think of the purposive unity the form of an organism exhibits; the fact that every part – in plants just as in humans – serves a function that can be understood only within the context of the functioning or maintenance of the whole; and where thus every part is at the same time equally both end and means. Think of the migrating bird again. Or of a tree:[167] Its roots stretching into the ground, "searching" for the nutrients and being the anchor; its leaves up high, "reaching" for the sun; and falling off come winter. – All of it, unless something goes *wrong*, so that the tree itself may live. All of it intelligible only by appealing to a function, to a *purposive activity* of the tree that aims at its "good" (life); and yet clearly without any consciousness or reason to speak of. But even we, as Nietzsche points out, often "perform – without any deliberative effort – the most purposive, 'logical' action"; as when our foot slips, for example.[168]

Unsurprisingly, it is here that Nietzsche thinks the notion of the drive (or instinct)[169] – as just such perfectly unconscious yet "intelligent" purposiveness – must enter the picture (which

---

166 For the original German text, as well as an English translation of the author, see "Appendix A" in Crawford C., *The Beginnings of Nietzsche's Theory of Language*, Walter de Gruyter, 1988.

167 This is the example Kant himself uses in §64 of the third *Critique*

168 D 133.

169 In "On the Origins of Language" Nietzsche praises Kant for discovering the "essence of the instinct" in "the wonderful antinomy that something purposeful can be without consciousness". (See Crawford C., *The Beginnings of Nietzsche's Theory of Langua*ge, p. 226.)

is why it is supposed to hold such immense explanatory power). As Nietzsche puts it: "Of all forms of intelligence discovered hitherto, 'instinct' is the most intelligent."[170] Indeed Nietzsche seems to believe that these unconscious, instinctual processes that go on beneath the surface already exhibit a form of inferential "reasoning"; and can thus serve as a basis for conscious discursive and linguistic activity. As such, insofar as we remain within the practically engaged, "intentional stance" at least, reality as such would have to be understood as conceptually structured. Accordingly, not just humans but also other animals would experience the world in a "categorially" articulated way – even if they do not *take* it as such (and thus cannot use language and so on).[171]

> For the longest time, conscious thinking was considered to be the only thinking: only now does the truth dawn upon us that the greatest part of our mental activity goes on unconsciously and unfelt by us.[172]

So we get to the big question. What do we do with self-consciousness and reason? That is, with our putative ability to not only *know* our purposes (and not just ours) but also to *set them*? Why should we posit something that seems to escape, even break from, this *instinctual* unconscious logic of nature? Can we not explain ourselves using the same natural principles?

We know what Nietzsche's answer is. Armed with the conception from above, Nietzsche apparently came to be convinced that we could now account for just about everything while

---

170 BGE 218. (See also GS 11.)

171 This distinction tracks quite closely Riccardi's distinction between *reflective* consciousness (*qua* self-consciousness), mediated by language and dependent on linguistic capacities, and *perceptual* consciousness, the kind of awareness and "knowledge" we can ascribe to non-linguistic animals. (See the sixth chapter in his *Nietzsche's Philosophical Psychology*.)

172 GS 333.

staying at the unconscious, bodily level – and reduce *self-consciousness* to nothing but "a mirror" in which we observe ourselves; an ability to know (or *think* we know) in what "state" we currently are (and to communicate that to others).[173] Already in Chapter II we saw that this solution has its problems. Most importantly, this self-consciousness, no matter how minor, remains still something quite distinct and – as Nietzsche himself at times (perhaps inconsistently) admits, certainly seems to have *a role to play*. Furthermore, when Nietzsche discusses this role, it seems to grow and expand, despite his explicit claims, into something that's anything *but* superficial – seeing as it's supposed to be *necessary* for our socialization, the development of language, and thus, it would seem logical, just about all fruits of civilization (but, alas, morality too).[174] Moreover, as others have notices,[175] Nietzsche's repeated insistence that consciousness is harmful and dangerous[176] seems like a claim that should be difficult to reconcile with the parallel assertion that it is inert.

Nietzsche's strategy here, consistent with his overall project of "replacing" conscious thought with unconscious instincts, is to claim that reflective thought itself is merely a *product* or *tool* of selected the drives. And this is the key. As I already remarked earlier when discussing the drives, what happens when when we try to account for why we act and think a certain way, for why we did this or that, we exclude these forces from our explanations (and thus subsequently from how we understand ourselves in general). We exclude them precisely because they are

---

173 Cf. GS 333.

174 See mainly GS 354.

175 For example Katsafanas in his *The Nietzschean Self*,

176 He even has a nice evolutionary tale to tell in support of his position in GS 11: "Consciousness is the last and latest development of the organic, and consequently also the most unfinished and weakest of these developments. Countless mistakes are produced by consciousness that cause an animal or man to break down sooner than necessary."

unconscious; that is, because reflective thought (as their product or by-product) cannot reach them. But that simply means that in rational reflection we become – by definition – *literally deluded* about the actual reasons or motives for our actions.

Thus, because the only type of motivation Nietzsche acknowledges, the only "drive", is that of the unconscious, instinctual *drives,* reason – rather than a servant with at least sort form of autonomy (such as that of selecting the correct means) – becomes *thoroughly enslaved*. Except, of course, it was never free in the first place. Nietzsche is quite explicit about this. For instance, at one point in *The Gay Science* he tells us that the intellect is "nothing but a certain behaviour of the drives towards each other".[177] And his discussion of what we experience as willing in *Beyond Good and Evil* suggest the same picture.[178] Yet the most telling passage is certainly the following one from *Daybreak*, where Nietzsche writes:

> "*That* one desires to combat the vehemence of a drive at all (...) does not stand within our own power; *nor does the choice of any particular method*; nor does the success or failure of this method. What is clearly the case is that in this entire procedure our intellect is only the *blind instrument* of another drive."[179]

To connect this to our original topic, it is not difficult to see how this might have some serious consequences for morality and altruism. If a moral act is one performed out of *selfless motives*, should we be searching for selfless unconscious drives? We might be able to appeal to biological notions like kin selection and "altruistic" behaviour entailed in[180] – but what does that have in

177 GS 333.

178 BGE 19. (See also KSA 12:26; 11:594–595.)

179 D 109 (emphasis mime).

180 And Nietzsche does indeed speak of our "herd instincts", which he sees as evolutionary products of mankind's gradual socialization (itself necessary because of our individual helplessness), responsible for the birth of both morality and reflection (which are for Nietzsche intimately connected). See, for example, GS 116, 328, 354.

common with how we understand selflessness; namely, with our *conscious recognition* of others as ends in themselves? If the ends of our actions are "given" – rather than being open to questioning and thus "up to us" – what hope is there for morality as traditionally conceived?[181] What hope is there for holding myself accountable?[182] For being answerable for *the reasons why I am doing what I'm doing* – and hence for reasons-responsiveness and normative force of arguments? There is an answer to this question. – *None*. On Nietzsche's picture, whether I come to believe something, judge it to be true (even something completely trivial, such as that the table in front of me is made out of wood), does not depend on my taking it to be the case. It does not depend on any objective normative warrants. Instead, it depends on whether the causally efficacious processes that go on behind our backs (or "under the surface",) just happen to trigger or "excite" the appropriate drive(s).

181 Indeed Nietzsche makes the link between morality and rational willing (or freedom) repeatedly explicit; as when, immediately after dismissing any morality that should be based on wholly selfless actions, he adds: "If only those actions are moral which are performed out of freedom of will, as another definition says, then there are likewise no moral actions!" (D 148)

182 For Nietzsche's sweeping denial of responsibility, see the fifth chapter of Leiter's *Moral Psychology with Nietzsche.*

> "Isn't this the 'terrible' truth: that no amount of knowledge about a deed is *ever* sufficient to bring it about, that up until now the bridge from knowledge to deed has not been built even once? Actions are *never* what they seem to be! (...) Moral actions are in truth 'something else' – we can say no more: and all actions are essentially unknown."[183]

Traditional morality believed that the essence and value of a deed derived from its intention. Nietzsche tells us that that "the decisive value of an action lies precisely in what is unintentional in it."[184] Morality of intentions and personal responsibility is consigned to the dustbin of metaphysical superstitions – one more consequence of *bad psychology*. Instead enter: the "extra-moral" age.[185]

It should be obvious by now that strategy of Nietzsche's has the potential to disrupt far more than selfish and selfless intentions; more than morality even. For it touches the very nature of ourselves – of our being self-conscious agents, responsive to reasons, capable of asking ourselves what really matters to us, capable of *deciding* what ends we should pursue and how we ought to go about them. For as we can see (most clearly in the *Daybreak* passage quoted above), Nietzsche wants to reduce reflection, self-consciousness, the thinking Ego, or whatever else we wish to call to a mere by-product, something whose source lies in "unreason": In in the bodily self understood as an arena of mutually competing drives and their specific *affective* "interests" from which all action must spring and from which it cannot stray. In the mysterious Self as the instinctive "great reason" that at all moments pulls us, unconsciously, towards purposes we are

183 D 116. (See also BGE 287.)
184 BGE §32.
185 *Ibid*.

not privy to and about which we can have no say. As a result, none of us are really subjects *of* our actions and desires. Instead, we am subject *to* them. But even this seems to be going too far. For *who* is this "I" that is subject to these things? Who is performing all these putatively purposive acts on which the conscious "I" is supposed to "supervene"?[186] Seeing as Nietzsche makes our answering of the question "What should I do with my life?" completely vacuous – given that it is always already decided "for us" – we aren't *really* individual persons at all. As such, by making subjective intentionality – the condition of the possibility of any unified experience, be it action, belief, or desire – secondary and *inert*, Nietzsche dissolves the Self completely.[187]

And this leads us to our final problem. For if we affirm what we've been through in this chapter, what are we left with? All meaningful normativity, all purposiveness, all freedom, all talk of reasons, all distinctions between good and bad, success and failure, power and weakness, health and sickness, life and death evaporate into air. But how could that be? How can we reconcile this with the "practical" side of Nietzsche's project (including the force his writings might have on the reader)? With all the talk of agents that "make promises", understand themselves and give seemingly rational justifications for their actions and decisions*?* How can

---

186 This brings back our homuncular worries from earlier; thought not so much with respect to the drives, but rather to the reflective, conscious "I". Nietzsche, as we saw, claims that these acts are the expressions of the unconscious, bodily, "creative" Self. However, because this self is essentially unconscious, and the intentional "I" its accidental by-product (something which may or may not accompany the Self's actions), can we truly call them acts? What exactly makes them distinct from mere events? How do we draw the line between something I do and something that happens to me? The only way is by having myself take myself to be intentionally doing something – undertaking a commitment by engaging in a norm-governed activity. But this is precisely what Nietzsche denies (or thinks only "apparent").

187 Sebastian Gardner correctly observes that Nietzsche's theoretical (anti-realist) account of the Self contains no trace of "how the self is for itself"; that is, no subjectivity. But when it comes to Selfhood or agency, subjectivity is of course a *sine qua non*! (See Gardner S., "Nietzsche, the Self, and the Disunity of Philosophical Reason," in Gemes K. & May S. (eds.), *Nietzsche on Freedom and Autonomy*, Oxford University Press, 2009, p. 11.)

Nietzsche, the philosopher of self-overcoming and self-creation, of the *Übermensch*, simultaneously deny that we are capable of *leading* our lives?

But of course we already have the answer. As we saw at the beginning of this chapter, Nietzsche doesn't just reject these things as non-existent. In a way, he doesn't even dismiss them as *unimportant*. On the contrary, he is perfectly ready to admit that from our lived, first-person perspective it matters very much to us that we be alive and that we see meaning in what we do and undergo.[188] Or rather – and this is the key – it *seems* that it does. It only seems so because, according to Nietzsche, agency, life, freedom, reason, etc. are not *constitutive* of what makes us us. Instead, they are merely accidental *fictions* or "user illusions" – "in themselves" perfectly superfluous, but "for us", *phenomenally*, they are everything.

Everything depends on how we solve this problem. After all, the main purpose of this entire work was to make explicit the necessary condition of agency. But now we face the possibility that agency as such simply isn't "real" – that what we really are is nothing but self-deluded puppets of blind, non-rational fate. Accordingly, our next task couldn't be clearer. We must examine Nietzsche's claim and see whether he can truly entitle himself to it. In the final chapter of this work I attempt to do just that.

188 In the very last section of the *Genealogy* Nietzsche writes: "The *senselessness* of suffering, *not* the suffering, was the curse that hither laid on mankind." (emphasis mine). And in the *Twilight* he famously tells us that he who has a *why* in life can bear almost any *how* (TI "Arrows and Epigrams" 4).

VII. The Disunity of Reason

On the face of it, it would seem that Nietzsche's position assumes a rather familiar, perhaps even intuitive, separation of the "subjective" and the "objective" – a division between *our knowledge* of the world and the world *as it really is*. To state it in more or less Kantian terms, Nietzsche seems to believe that because we are finite, fallible, material beings, the way the world appears to us is made possible only by certain "transcendental" conditions that we ourselves "provide". But – given that it is *us* who provide them, because our "access" to the world is always necessarily mediated (which is to say: distorted or falsified) by these conditions – the resulting knowledge must be understood not in terms of some naive realism, but rather as mere *appearance*. In Nietzsche's own words: Thought is nothing but "*an interpretation according to a scheme that we cannot throw off*".[189] One can almost imaging finding these words on the pages of the first *Critique*. In any case, the important point is that for Nietzsche, as Béatrice Han-Pile observes, "experience (...) has what looks like transcendental conditions in that it is necessarily structured by '*a priori* forms'."[190]

But of course, experience is not just experience of the external world. Nietzsche never tires of pointing out that our "*inner* world is also an 'appearance'!"[191] Which is to say that our own self-understanding is itself structured by these "*a priori* forms" of thought, or a specific "conceptual scheme" – that we cannot simply discard to understand ourselves differently. We must assume them, as we simply *couldn't live without them*.[192] Which is to say, any effort to deny

189 WP 522.

190 Han-Pile, B., "Transcendental Aspects, Ontological Commitments, and Naturalistic Elements in Nietzsche's Thought," in Gardner S., & Grist, M. (eds.), *The Transcendental Turn*, Oxford University Press, 2015, p. 198.

191 WP 476

192 See for example BGE 4, 11; WP 493, 515.

them would point to *practical unintelligibility*. Accordingly, it is also most evident when applied to the concept of agency itself. In particular, the notions of freedom and the Self. Which is, of course, precisely what we are now scrutinizing. Let us consider freedom first.

The question of "Nietzsche's Freedoms"[193] has been a long-standing exegetical puzzle. On one hand, Nietzsche speaks of intentional and purposive agents that make decisions, engage in arguments, "make promises", resist (or fall prey to) the complicity of custom, set their own ends, create their own laws, and strive to become who they are. As we have seen throughout this work, this puts a much needed emphasis on the fact that freedom is, first and foremost, about *agents* who *express themselves* by articulating and negotiating what it is that matters to them. On the other hand, however, Nietzsche seems committed to seemingly "overriding" metaphysical determinism or fatalism.[194] A fatalism that often appeals precisely to our internal constitution (the drives) over which, as we have seen, we have no control. The main question thus asks: How is this possible? How do we reconcile such contradictory positions?

I believe the answer rather simple in fact; and it has already been given: Nietzsche is a sceptical, anti-rationalistic subjectivist who thinks that the way we must take things when practically engaged is not at all "aligned" with how things really are. Or perhaps: the way we must take things is itself internally contradictory (seeing as consciousness is merely a limited and faulty tool). After all, this is precisely how it is for Kant: It is *because* nature is appearance – *because* it is intelligible to us only in relation to us (as structured by the *a priori* "forms" we provide); that is, as *phenomenal* – that we must understand it as deterministic. (Given that one

193 See the volume edited by Gemes and May, *Nietzsche on Freedom and Autonomy* (and especially essays 1, 2, 4, 6, 7, and 8).

194 See HH I.39; D 109, 124, 130; GS 108; BGE 22, 231; TI "Morality as Anti-Nature" 3, 6; or WP 458 (among others).

of the categorial requirements rationally necessitates that we grasp all temporal unfolding in accordance to the necessary causal law of succession.)[195] And yet, as Kant wants to tell us in the second *Critique*, we must also understand ourselves as spontaneous and free (or "practically rational"). We simply couldn't do otherwise.

Now, Nietzsche has very little patience with the noumenal as well as with "intelligible freedom". However, he is clearly aware of the practical unintelligibility of denying freedom. For when he affirms his uncompromising fatalism, he states it cannot really have any practical consequences. For see, once we accept it, the notion of trying to resist one's fate – in *any* way – simply ceases to make sense. It becomes, from any practical standpoint, completely unintelligible. For suppose that you come to be convinced that purposive, self-determining agency is an illusion and that we truly are just puppets on strings. How could you respond to something like this? How could you acknowledge it in your life? Are you just supposed to sit and wait until something determines you from the outside? But clearly, from your lived, intentional, first-person perspective that sitting and waiting would also be *your doing*. One simply cannot take oneself as not deciding, or doing "nothing"; any such "nothing" will always be *something*. As Nietzsche himself noted, when one affirms fatalism, the struggle against one's fate becomes imaginary *but so does the resignation.*[196] Thus, nothing really changes.[197] Of course,

---

195 See the "Second Analogy" section in the *Critique of Pure Reason*.

196 HH II.61.

197 One may think that while we cannot, in practice, take ourselves as determined, we can take this into consideration in matter such us how we approach responsibility or punishment. But of course if determinism is true, then it is *all the way down*. Our re-*acting* to it has always already been decided and is thus purely mechanical and passive. The notion that it is *us* adjusting to anything (our acknowledging of the absurdity of punishment, for example) is, once again, to fall prey to the illusion. There simply is no acting, no deciding, no acknowledging. – And yet, "for us", this is all there is.

this does make one wonder whether such a doctrine isn't completely vacuous. A pragmatist, for example, should surely dismiss it as completely inconsequential – and thus untrue.

The situation is exactly alike when it comes to the Self. As Sebastian Gardner points out, because Nietzsche's anti-realist conception of the Self is purely "external" (pays no attention to first-person subjectivity), it generates a "discrepancy with the (necessarily internal) practical point of view"[198] – which seems to presuppose it. And in fact, at certain points Nietzsche himself seems to acknowledge that the "I" *is* in fact one of those transcendental conditions that we cannot "throw off".[199] In WP 487, for example, he describes "*our belief in the 'I'* as a substance" as one of the "presuppositions on which *reasoning* depends".[200] But then, surely, any attempt to explain the I away will inevitably end up assuming what it wants to reject, and will thus positively contradict itself. Or can reason somehow circumvent its own constitutive conditions? Or are we to reject the Self (and reason) by some other means? – This is the strange puzzle that we must now turn to. For if it can be proved that the reality of self-consciousness and rational agency *cannot* be challenged this way, then Nietzsche's sweeping sceptical attack would be once again repelled as a whole.

Firs of all, what should we make of this discrepancy? Gardner argues that such basic tension ultimately points to the fact that Nietzsche never quite managed to reconcile the practical (and ultimately, transcendental) aspect of his philosophy with the theoretical (naturalistic) one; and that what we get as a result is merely a disappointing Humean dualism: on one hand, we have "the non-existence of a self for theoretical reason", and on the other "the necessity of the

198 Gardner S., "Nietzsche, the Self, and the Disunity of Philosophical Reason," in Gemes K. & May S. (eds.), *Nietzsche on Freedom and Autonomy*, Oxford University Press, 2009, p. 11.

199 See *Ibid.*, p 5.

200 WP 487.

assumption of a self for *practical* purposes".[201] In other words, Gardner suggests that there comes a point when Nietzsche simply throws his hands up and declares that it is simply our lot in life – being the amphibian creatures that we are, inhabitants of both the natural world and the spiritual world – to never achieve a proper unification, a fully integrated world-view and self-conception.[202] Our "philosophical representations" simply cannot do justice to "the purposes of life".[203]

On one hand, this strikes me as an accurate diagnosis. Nietzsche did see that his project of "translating man into nature" had its limits. His admonition of the Stoics for pretending to "live according to nature" (which Nietzsche declares impossible) is an instructive example of this.[204] Yet there is more that needs to be said. Most importantly, while Nietzsche may not have seen how to reconcile the two standpoints, he is usually very adamant about which one must go first (albeit, once again, merely in theory). In fact, sometimes he goes so far as to call the other perspective a fiction. Consider, for example, Nietzsche's claim that while "in 'thinking', the ego is presupposed", that alone proves nothing – for "a belief can be a condition of life and *still* be *false*".[205] Or, in fact, his early popular (albeit unpublished) essay *On Truth and Lying in Non-Moral Sense* – one of the main theses of which is that even "true" or factual knowledge (and thus

---

201 *Ibid*., p. 16. Gardner rightly notes (and one can almost hear Thomas Reid uttering those words) that: "Disunity of reason in the natural consciousness of the Humean subject never surfaces for the subject itself – it comes to light only in Hume's study, and Hume's reflections in his study are forgotten as soon as he steps outside; so the Hume who knows that he has no self and that his reason is disunited, never gets to meet the Hume who thinks he has a self and whose reason is disunited." – This is not only accurate as a description (of both Hume *and* Nietzsche), but it should also make us highly suspicious that their theoretical accounts are simply inadequate.

202 *Ibid*., p. 28.

203 *Ibid*., p. 29.

204 BGE 9. (Of course, the problem might very well be with the specifics of the project/naturalism itself.)

205 WP 483. (For some other examples of Nietzsche's "affirmation" of the necessity of certain *a priori* "constitutive rules" without which we couldn't think and live – but which he still considers false – see BGE 4, 11; or WP 477, 482, 489, 493, 512.)

all knowledge) is really a "lie" because it is *necessarily* a subjective, anthropomorphic distortion or falsification of the inaccessible and unknowable objective "X" of nature[206] – a "user illusion" we cannot but suffer, for it is an inherent, inevitable part of what it means to be human, all too human: a being whose *sole means* of accessing the world (the external one as well as the inner one) is that of a simplistic species-specific, and to some degree individual-specific,[207] *conceptual scheme* that has absolutely nothing in common with reality as it really is.[208]

But how could that be? Does Nietzsche think we can somehow *go around* assumptions he himself deems "necessary" for the thinking (or living) as such? How? Wouldn't this "going around" itself involve thinking – and thus those very assumptions or categories? What sense can we make of the notion of a "necessary illusion"?[209] Or of truth being nothing but an "error"[210] and "a mobile army of metaphors, metonyms, anthropomorphisms"?[211] *None*. After all, do we have access to some perspective "outside" knowledge"? Or "beyond thought"? Is there supposed to be some "purer" experience of the world that goes beyond attempts at comprehension? If

---

206 See TL pp. 144–5 (in particular, Nietzsche's claim that "nature knows neither forms nor concepts (...) but only an 'X' which is inaccessible to us and indefinable by us").

207 Which is to say, we are each different, yet (presumably because of heredity and evolution) *similar enough*.

208 As Nietzsche remarks, how could there ever be a metaphysical correspondence between nature as it really is and our knowledge of it? What does reality have in common with what is first a nervous stimulus, then a conceptualized, perceptual image, and finally an articulated sound?

209 For instance, Nietzsche's repeated insistence that concepts, like forms, falsify because they abstract from individuality and complexity and produce unities and equalities where there are none. As he puts it, no two leaves are perfectly alike after all (TL p. 145). Yet it is easy to see that without concepts there would be nothing determinate to be known, and that it is only *via* concepts that we can investigate the details. The fact that sometimes it proves useful to generalize is hardly an objection against conceptual or discursive knowledge. Moreover, Nietzsche's admonition of logic for projecting equalities upon a formless, chaotic word of becoming cannot stand either. The concept of difference makes sense only next to identity. And once we affirm that something is more or less *alike*, then we presuppose some criterion in light of which the "more alike" is *more alike*. Utterly undifferentiated, chaotic, incoherent world (see GS 109; WP 569) is no world at all.

210 See WP 493: "*'Truth' is the kind of error* without which a particular kind of creature could not live."

211 TL p. 146.

there is, we certainly cannot say anything about it; just like we cannot say anything about the "mysterious", "undefinable" "X" – which is in any event just Nietzsche's own name for the unknowable *Ding an sich*.[212] But if we can say nothing about them, how would we even know they are there (and that they ought to count *as knowledge*)? What role could they possibly have in our philosophical accounts? To be sure, we can give these empty, abstract "ineffables" *names*, but names are hardly helpful. What we need are *determinations*: relations, properties that exclude other properties; something that can help us situate the notion via inferential implications and incompatibilities, give it specific *content*, make it something distinct from other such possible "X's".[213] If we cannot do that – if the concept isn't part of the normative space of reasons – then it isn't a real concept at all, but merely a vacuous product of our own thought. Non-sense – and thus: *no-thing*.[214]

Of course, Nietzsche himself recognizes that we cannot know anything but this world; and that we cannot get "outside" our consciousness, "outside" thinking. But then, how could all our

---

212 Commentators are often opposed to the idea that Nietzsche kept holding fast to the notion of the "thing in itself" even after his early Schopenhaurian days. And it is true that from *Human, All Too Human* onward Nietzsche gets progressively more and more dismissive whenever he brings up the notion. In HH I.9 he says that while we can hardly cannot prove the metaphysical world does not exist, its potential existence should remain utterly inconsequential to us. Later, especially in GS 374, he starts emphasising his relativistic "perspectivism" (the fact that *some* perspective is always necessary if there is to be knowledge, even though these perspectives might be completely unlike ours). And in BGE 16 he even call the *Ding an sich* a "*contradicto in adjecto*". And yet, despite all his dismissals, it would be a mistake to take Nietzsche at his word. His "epistemology" – because it is a kind of subjectivistic psychologism – remains committed to this dualism all the way to the end, even if unknowingly (although at times it does seems to slide into radical relativism).

213 Again, surely there must be something that determines both of these "realities" (or types of knowledge) *as two instances of what we ought to count as reality* (or knowledge). If there isn't, then such inaccessible, original *Urgrund* will be just another ineffable (or name for being), completely indistinguishable from all others – "the night in which all the cows are black", as Hegel aptly called it.

214 Nietzsche says: "Ultimately, man finds in things no more than what he himself has introduced into them." (WP 606). What a surprise then that when we abstract from every determinacy and every means we could ever have of cognizing the "X" we end up with something unknowable: a mere "beyond", an empty, worthless residue.

knowledge be false? For one, if all knowing necessarily distorts "the facts", then why listen to anything Nietzsche has to say? How could he know that consciousness is superficial? How could he come to know what "really moves us"? How could he ever say anything *about* anything? It seems Nietzsche trapped himself in an incomprehensible position. One one hand, he recognizes that we cannot see "around the corner";[215] that we do not have access to anything but *this* world and by no means other than *our* means. But on the other, he also wants to say that the *practical* and transcendental necessity of affirming agency, the Self, freedom, rationality, etc., *can* in fact be done away with within some purely *theoretical,* "scientific" or naturalistic account of what these concepts amount to (that is, *not* from our subjective perspectives, but *objectively*).

But this just cannot be. Giving a theoretical account is a form of practical – subjective, intentional – activity. We cannot just decide to set the "conceptual scheme" aside for a minute. No, as Nietzsche himself repeatedly affirms (and confirms by his own writings and doings), *we cannot throw it off*.[216] In absolutely *everything* we ever do or think, our self-conscious conceptual activity – determination of what is the case via conceptual or "categorial" articulation – is always at work. That means, however, that rational agency, selfhood, and freedom cannot be merely "pragmatic" notions or "useful fictions". They are, plain and simple, the metaphysical conditions of intelligibility of any practice as well as theory – including the sceptical position itself.[217] But then – how could they "*still* be *false*"?[218] First, it is trivial to see that the notions of

---

215 Cf. GS 374.

216 WP 522.

217 Why do we need conceptual coherence for practices? Because for a practice to be the determinate practice it is, it needs to be governed by specific (albeit often implicit) norms; and because for some practice to even be *practice* (rather than a set of mere events), it must embody purposive intentions. – All of which is unintelligible without conceptual determination.

218 WP 483.

falsehood and falsification become intelligible only as internally related to their dialectical counterparts.[219] And second, appealing to some mysterious X "in itself" was already dismissed as empty and incomprehensible mystifying.[220] Thus, to conclude: *they couldn't*.

Once again, we witness Nietzsche rightly criticising what is a philosophically credulous position (naïve realism) – but then going way, way to far in the opposite direction. Specifically, he knows that we don't simply look and absorb "the facts" (in which case no science or philosophy would ever be needed), but rather that all knowing or is always judging and *interpreting*. Yet he misses that the activity of interpretation necessarily presupposes that there is a way how the object *ought to be* judged; that in interpreting one is – necessarily – trying to get something *right*. If there wasn't, there would be no meaning at all; no determinate content (as the "*what*" of the *interpretandus*); nothing our interpretation could be *about*.[221] Every judgment would be absolutely arbitrary, formless, indeterminate. Thus, whenever one judges a certain interpretation *false*, one is at the same time implicitly assuming a possibility of getting it *right*. Or, put differently, it is because there is some way things ought to be understood that we can also

219 Similarly, how something merely *seems* or *appears* is necessarily "parasitic" on how it objectively *is* (that is, the moment of objectivity – how something ought to be taken – is already "built into" any judgment).

220 The decisive to ask oneself here is: What could it *possibly mean to be* other than *to be intelligible*? Just consider what we actually mean when we say that we know what something *truly is*. We mean that we have achieved such comprehension of the subject matter that it leaves no conceptual or empirical remainder behind. That the thing in question became entirely intelligible and *known*.

221 Another way of putting this point is by noting that in saying anything meaningful at all (i.e. in judging anything to be such and such) one is implicitly committing himself to the principle of non-contradiction (and to internal coherence or grounding of one's claims). Nietzsche, predictably, denies the principle as another anthropomorphism (see WP 516; in general, he thinks that the rules of logic represent the "fundamental falsification of all events", as in WP 512, 517, or BGE 4), but of course, if he truly believed it, he couldn't ever say *anything*.

be mistaken.[222] Accordingly, Nietzsche's principled scepticism cannot even account for itself – for its own determinacy, its own distinctness as a contentful *philosophical position*.

Where do we go from here? Can Nietzsche somehow untangle himself from this large contradictions he got himself into? His strategy consists in challenging the claim that the constitutive *a priori* forms of thought (such as the 'I' or freedom) aren't *really* universal[223] – and so could, potentially, be bypassed. Nietzsche seems think of them in quasi-transcendental, putatively "naturalistic" terms. Namely, *pragmatic, evolutionary* ones. Consequently, he tries to appeals to contingent, provisional "necessity" of a "merely" *psychological* kind – that is, to the idea that we must *believe* them to be true (and structure our practices around them) for the sake of our own preservation and growth.[224] As Nietzsche writes in the *Nachlass:*

> "I regard the most strongly believed *a priori* "truths" as – *provisional assumptions* (...) But are they thus truths? What a conclusion! As if the preservation of man was a proof of truth!"[225]

Han-Pile rightly notes that Nietzsche here anticipates what Foucault would later come to call "historical *a priori*": constraints that are both practically necessary for us – in that we, within our

---

222 In TI "Reason in Philosophy" 2, Nietzsche asserts that: "Reason makes us falsify the testimony of the senses." This might be read as implying (especially along the following section 3), that the sense are somehow pure and undistorted ("true"). But this is the exact same problem. As Descartes' famous example with wax demonstrates, the sense themselves are utterly inert without thought – without judgment that must saturate them and conceptually articulate what it is we "sense". As Kant remark in the first *Critique*, the senses do no lie, to be sure, yet not because they always tell the truth – but rather because *they do not judge at all* (A923/B350).

223 Of course even for Kant the *a priori* forms aren't completely universal in that there could be other kinds of beings that are not constrained by "discursivity". They are, however, universal *for us*. Nietzsche wants to weaken even that claim. But, in fact, already Kant's position proves too subjectivistic. For if there is something like knowing or rationality, than stipulating different forms *of* rationality implies there must be something they have in common (and if there is, then we can know it and "translate" it).

224 BGE 11. (See also BGE 4; WP 515.)

225 WP 497.

historically specific set of practices, couldn't very well think and act without them – but also historically contingent and variable.[226]

Now, of course there are a great many historically variable concepts that are so deeply embedded as to seem absolutely necessary and universal. Similarly, nobody disputes that our non-empirical, a priori conditions on knowledge are themselves dependent on empirical circumstances and have been historically acquired. And yet, Nietzsche himself doesn't give us any coherent suggestions on how something like agency or subjectivity might be bypassed. His attempts at explaining the "I" away via contingent factors, such as the grammatical (subject-predicate) structure of our languages,[227] cannot do the work he needs because – once again – they invariably presuppose what he wants to disavow. In this particular case, the fact that language requires a speaker; an individual capable of being initiated into its mysteries – and that means: rational, reasons-responsive, self-conscious agents.

And equally with freedom. We might want to say that that our conception of ourselves as free is merely a historical result of the will to power at work (of particular power relations in given societies or something along those lines). Yet how could power ever be *power* if there were no free subjects over which it could be exercised? That is, purposive agents whose self-determination may be externally influenced and impeded? Accordingly, while what counts as freedom certainly isn't set in stone and remains dependent on historical practices, agency *as such* – as a constitutive feature of certain kinds of beings (beings that are capable of determining what, specifically, should count as freedom) – does not. No "historical a priori" can ever explain

226 Han-Pile, B., "Transcendental Aspects, Ontological Commitments, and Naturalistic Elements in Nietzsche's Thought," p. 224.

227 Cf. TI "Reason in Philosophy" 5; WP 522, 551, 562.

or substitute for a first-order philosophical account of what it means to be an agent. Quite to the contrary. What we must ask is this: What is agency – such that it must be historically constituted? *And such that we cannot avoid it?*

And there is only one way to truly answer that question. We must try to give an account of what it could possibly mean to be an agent. That, however, can only be accomplished by taking reason seriously – and by having it scrutinize and critique itself. Accordingly, we simply cannot afford to take our psychological or empirical assumptions for granted. Nietzsche, as so many others, has become so convinced of certain "naturalistic truths" (such as the illusoriness of freedom, rationality, the Self, etc.) that he simply couldn't entertain what he thought incompatible – even if it proved *rationally necessitated*. Yet as we just saw, even the disowning of one's reason and freedom cannot be understood as anything other than (mistaken) expressions of a *rational, self-conscious agent* – an agent who takes himself to have reasons for his beliefs and actions; and who is at all times striving to be the person he thinks that he should be.

Conclusion

At long last, we have arrived and the end of this work. As we have seen, in the end, Nietzsche's pervasive subjectivistic anti-rationalism fell prey to nothing other than reason itself. It became clear that we cannot escape, reject, or depreciate reason (with all that it necessarily entails) because any attempt to do so will, once again, necessarily involve *reason* and all that it entails: most importantly, openness to truth (as a distinct internal *norm*), and of course, self-determining, purposive agency itself. Nietzsche's "anti-rationalism" thus turned out to be nothing but a form of *irrationalism.* The very highest form, in fact. For what Nietzsche commits himself to is not only an inconsistent, self-contradictory position, incapable of entitling itself to its own claims, of accounting for its own unexamined premises (which, if examined, end up contradicting what the position explicitly avows) – but *self-conscious reason's denigration of its very own self*.

Accordingly, Nietzsche's greatest error lies in nothing else than his limited conception of reason as such. A conception that grasps reason in *merely* pragmatic, instrumental terms – that is, takes reason itself *as an instrument.* Yet, repeatedly, we have seen that an instrument – or *means* – is what reason simply cannot be. Every recognition of something as "useful" or "adaptive", for example, already presupposes it. We cannot outrun reason; or step outside of it. We cannot pick it out as means towards something else. But then – we must simply affirm it as end. An end? Indeed. But not just any end. For reason isn't some wholly separate, inert faculty for purely "factual belief" – as Nietzsche himself rightly emphasises. In fact, it is among Nietzsche's greatest merits that he targets and challenges those who regard reason as divorced

from life and its purposive striving, from the valenced, desiring responsiveness to instruments and obstacles is an empty abstraction perpetuated by contemporary. But reason also something insignificant added "on top"; some "extra bit" or "layer". No. It is the *form* of a particular kind of animal; of a particular kind of internally purposive, embodied agent that simply couldn't be itself without it. Indeed, when Nietzsche says the purposes of reason are to be found in life, he doesn't know how right he is. For in our case, life and reason are *one*. That is, rationality is nothing but a distinctive actualization of biological autonomy; the highest, freest form of organic autopoiesis. As such, the purpose of rational life can be nothing other than to strive to constitute and maintain itself as such: As the self-knowing, self-constituting activity of living characteristic of embodied, desiring, rational agents, capable of setting their own ends. Agents that strive towards the *good* (what ought to be done) *by trying to get things right* (the true) – by being responsive to reasons as objective warrants for their beliefs and to the wordly means to their ends. Agents that hold themselves to their commitments – *to who they ought to be, and what it is that ought to matter*.

Bibliography

Nietzsche's Works

Sämtliche Werke, Kritische Studienausgabe in 15 Bänden, De Gruyter, 1980.

The Birth of Tragedy (1872), in *The Birth of Tragedy and Other Writings*, Cambridge University Press, 1999.

On Truth and Lying in Non-Moral Sense (1873), in *The Birth and Tragedy and Other Writings*, Cambridge University Press, 1999.

Human, All Too Human (1878/1880), Cambridge University Press, 1996.

Daybreak (1881), Cambridge University Press, 1997.

The Gay Science (1882/1887), Cambridge University Press, 2001.

Thus Spoke Zarathustra (1883/1885), Cambridge University Press, 2006.

Beyond Good and Evil (1886), Cambridge University Press, 2002.

On the Genealogy of Morality (1887), in *On the Genealogy of Morality and Other Writings*, Cambridge University Press, 2017.

The Twilight of the Idols (1888), in *The Anti-Christ, Ecce Homo, Twilight of the Idols and Other Writings*, Cambridge University Press, 2005.

The Anti-Christ (1888), in *The Anti-Christ, Ecce Homo, Twilight of the Idols and Other Writings*, Cambridge University Press, 2005.

The Will to Power, Penguin Classics, 2017.

Other Works

Bernstein J.M., *Torture and Dignity: An Essay on Moral Injury*, University of Chicago Press, 2020.

Clark M., *Nietzsche on Truth and Philosophy*, Cambridge University Press, 1990.

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

Constâncio J. & Branco M.J.M. & Ryan B. (eds.), *Nietzsche and the Problem of Subjectivity*, De Gruyter, 2015.

Crawford C., *The Beginnings of Nietzsche's Theory of Language*, Walter de Gruyter, 1988.

Darwin C., *The Descent of Man, and Selection in Relation to Sex* (1871), Princeton University Press, 1981.

Gardner S. & Grist M. (eds.), *The Transcendental Turn*, Oxford University Press, 2015.

Gemes K. & May S. (eds.), *Nietzsche on Freedom and Autonomy*, Oxford University Press, 2009.

Griffith M. & Levy N. & Timpe K. (eds.), *Routledge Companion to Free Will*, Routledge, 2017.

Hägglund M., *This Life: Secular Faith and Spiritual Freedom*, Penguin Random House, 2020.

Hegel G.W.F., *Lectures on the History of Philosophy* (1825–1826), Volume III, University of California Press, 1990.

Heidegger M., *Nietzsche*, Volumes III–IV, HarperCollins, 1991.

Hume D., *A Treatise of Human Nature*, Clarendon Press, 2007.

Hume D., *An Enquiry Concerning the Principles of Morals*, Hackett Publishing Company, 1983.

Katsafanas P., *The Nietzschean Self: Moral Psychology, Agency, and the Unconscious*, Oxford University Press, 2016.

Kant I., *Critique of Pure Reason*, Hackett Publishing Company, 1996.

Kant I., *Critique of Judgement*, Hackett Publishing Company, 1987.

Kronz F.M. & Tiehen J.T., "Emergence and Quantum Mechanics," in *Philosophy of Science*, 69(2), 2002, pp. 324–47.

Leiter B., *Moral Psychology with Nietzsche*, Oxford University Press, 2019.

MacIntyre A., *After Virtue: A Study in Moral Theory*, University of Notre Dame Press, 2007.

May S. (ed.), *Nietzsche's On the Genealogy of Morality: A Critical Guide*, Cambridge University Press, 2011.

McDowell J., *Mind and World*, Harvard University Press, 1996.

Poellner P., *Nietzsche and Metaphysics*, Oxford University Press, 1995.

Riccardi M., *Nietzsche's Philosophical Psychology*, Oxford University Press, 2021.

Riccardi M., "Nietzsche's Pluralism about Consciousness," in *British Journal for the History of Philosophy*, 24(1), 2016, pp. 132–54.

Riccardi M., "Inner Opacity: Nietzsche on Introspection and Agency," in *Inquiry: An Interdisciplinary Journal of Philosophy*, 58(3), 2015, pp. 221–43.

Richardson J., *Nietzsche's System*, Oxford University Press, 1996.

Richardson J., *Nietzsche's New Darwinism*, Oxford University Press, 2004.

Richardson J. & Gemes K. (eds.), *The Oxford Handbook of Nietzsche*, Oxford University Press, 2013.

Spinoza B., *The Collected Works of Spinoza,* Volume I, Princeton University Press, 1985.