



Open Research Online

Citation

Dries, M. (2015). Freedom, resistance, agency. In: Dries, M. and Kail, P. J. E. eds. Nietzsche on Mind and Nature. Oxford: Oxford University Press, pp. 142–162.

URL

<https://oro.open.ac.uk/43312/>

License

(CC-BY-NC-ND 4.0)Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

7

Freedom, Resistance, Agency

Manuel Dries

When we encounter a *resistance* and have to give in, we feel *unfree*, when we do not give in but compel it to give in to us, *free*. I.e., it is this *feeling of our increase of force*, which we name ‘freedom of the will’: the conscious awareness that our force *compels*, in relation to a force that is compelled.

(NL 1885, KSA 11, 34[250])¹

I can handle myself in the same way as a gardener his plants: I can distance motives from myself, in distancing myself from a place and company [*Gesellschaft*], I can place motives in my proximity. I can cultivate the propensity [*den Hang*], to proceed against myself in this gardener-like way, artificially [*künstlich*] or let it wither away.

(NL 1880, KSA 9, 7[30])

1 Introduction

The aim of this article is to show (1) that freedom and agency are among Nietzsche’s central concerns, (2) that his much-discussed interest in power in fact originates in a first-person account of freedom, and (3) that this novel understanding of the phenomenon of freedom informs his ‘theory’ of agency. I will argue that while Nietzsche questions the weight philosophers have given to the first-person perspective and consciousness, these remain essential not only to his initial analysis but also to his later conception of freedom and agency. While his rejection of metaphysical free will and moral desert has had a significant impact on contemporary ethics, the sense in which Nietzsche continues to use the term ‘freedom’ affirmatively remains largely unnoticed. He develops a sophisticated drive-driven psychological

¹ Throughout I use standard abbreviations for Nietzsche’s works. Nietzsche’s posthumously published writings (NL) are cited by year, KSA (Nietzsche 1988) volume number, followed by notebook and fragment number. For NL 1885, KSA 11, 34[250] I also use division IX of the KGW (Nietzsche 2001–), which offers Nietzsche’s late notebooks in diplomatic transcriptions that reveal his revisions, additions, and cancellations. I have relied on, and at times modified, existing translations of Nietzsche’s texts.

motivational account: reflective judgement and reasons can motivate by means of the affects or affective orientations agents have due to their drives; he claims that due to a strong preference—we could say with Mele, a ‘standing desire’ for freedom (it will soon become clear what Nietzsche means by ‘freedom’)—agents can generate the necessary motivational affects to unify their drives in view of (certain) long-term goals. Thus, when in his later philosophy Nietzsche envisages free agents who not only feel free but whose belief in their agency is justified, he has replaced the metaphysical picture (of agents who are mysterious, noumenally free, *causa sui* agents) with a naturalized, drive-driven psychological view of agency that he thinks has the resources to cope with the problem of affirmation that arises under nihilism conditions.

Let me clarify some terminology first. ‘Drive’ and ‘affect’ are key concepts for Nietzsche.² He uses ‘drives’ in the sense of relatively fixed and recurring tendencies that orient and structure behaviour and perception. ‘Affects’ refer to the first-personal experience of an orientation, usually expressive of a nuanced for-or-against attitude towards some feature of the world. Affects and drives are internally related: affects depend on and are expressive of an agent’s or self-system’s drives, thereby providing the qualitative structuring of an agent’s world.³ I will sometimes speak of dynamic, sentient self-systems rather than selves, persons, or agents. This is for heuristic reasons when I try to explicate the psycho-physiological mechanisms that Nietzsche discusses. ‘Dynamic’ emphasizes the relational and processual nature of a self-system that is not static or fixed and displays a capacity for self-regulation. ‘Sentient’ indicates that the self-system is not to be conceived as a physical system only, but rather, in contrast to a mere physical system, as capable of making sense of its environment, of having first-person experiential states, and of forming beliefs and goals that affect its behaviour. Finally, ‘system’ circumscribes merely a functional unit, a composite of drives, affects, habits, memory, language, beliefs, and—ultimately—also what Nietzsche calls ‘ideals’.

In Sections 2–4 I show that Nietzsche locates the basic idea of freedom in the experience of, or ‘what it is like’, to overcome resistances. In Sections 5–7 I sketch the Nietzschean hypothesis of a sophisticated, non-reductive motivational theory: due to an embodied, standing sense of self-efficacy, and drive (or standing desire) *for* self-efficacy (what Nietzsche refers to, rather obscurely, in his notion of ‘will to power’) agents generate, in unconscious and conscious mental simulations, the motivational affective states necessary for action. While Nietzsche often assumes

² See, among others, Richardson (2004, 2008), Janaway (2007, 2012), Leiter (2008), Gemes (2013), Katsafanas (this volume), Kail (this volume).

³ The relationship between affects, drives, instincts, and values is complicated and far from fully worked out by Nietzsche. Non-cognitivist and cognitivist elements feature in Nietzsche’s attempts to clarify their relationships.

that this motivational mechanism operates largely unreflectively—particularly in the ascetic he criticizes—it is argued that reflective judgements and conscious reasons may motivate via this embodied sense of self-efficacy.

2 Neither a Free nor an Unfree Will

Commentators are in agreement that Nietzsche devalues consciousness and leaves little room for any conventional understanding of free will and agency. He is unambiguous in regarding any conception of freedom as absolutely self-causing a metaphysical megalomania, an anti-natural invention typical of the Judaeo-Christian tradition. Free will thus understood would be incompatible with an understanding of human beings as natural entities. He thinks that the idea of free will was invented in order to warrant the responsibility needed for the purposes of socialization and the justification of punishment.

Commentators have suspected that Nietzsche's rejection of free will and his emphasis on psycho-physiological factors (such as drives and affects) over learned social factors or rationality lead him to reject the ontological domain of consciousness and the mental in favour of the ontological domain of the physical, thus subscribing to a version of determinism. This account is problematic, as Nietzsche explicitly rebuffs the simple inversion from a fictitious absolute autonomy to absolute heteronomy. He actually advocates giving up not only the belief in free will but also the belief in its opposite, what he calls 'the myth of the unfree will'. He proposes a change of categories⁴ in the much-discussed passage of BGE 21.⁵ Rather than conceiving of a self as either free or determined, or a compatibilism that allows the coexistence of moral responsibility and determinism, Nietzsche considers that the self is better characterized assuming degrees of freedom, degrees of strength and weakness. We see how puzzling Nietzsche's view can at first seem when we consider a note where he says that while willing is perfectly possible, this does not entail belief in the existence of will (see NL 1884, KSA 11, 26[254]). And yet, far from abandoning talk of freedom, the late Nietzsche still uses the term 'freedom'. As I will show, he refers first and foremost to a first-person experience of agency. Once we understand Nietzsche's re-description of the phenomenon of free will we understand the sense in which he rejects traditional free will and advocates an entirely different kind of freedom (a kind

⁴ On this passage, see Owen and Ridley (2003: 74) and Leiter (2007); also HH I 16. See also the earlier 1884 unpublished note: 'Das Problem von Freiheit und Unfreiheit des Willens gehört in die Vorhöfe der Philosophie—für mich gibt es keinen Willen. Daß der Glaube an den Willen notwendig ist, um zu wollen—ist Unsinn' (NL 1884, KSA 11, 26[254]).

⁵ 'If any one should find out in this manner the crass stupidity of the celebrated conception of "free will" / and put it out of his head altogether, I beg of him to carry his "enlightenment" a step further, and also put out of his head the contrary of this monstrous conception of "free will": I mean "unfree will," which is tantamount to a misuse of cause and effect. One should not wrongly materialize "cause" and "effect," as the natural philosophers do [...] The "un-free will" is mythology; in real life it is only a question of strong and weak wills' (BGE 21).

of agency par excellence, which enables self-control and the ability to set oneself and affirm complex goals and projects that are no longer ascetic).

Nietzsche is a naturalist of sorts who aims to 'retranslate man back into nature' and to shift the focus towards the body. However, he does not attempt anything logically impossible, namely to translate the mental into the physical, conceived as a different ontological realm that would, by definition, exclude the mental. His 'reduction' is really a 'translation', and it is explanatory rather than metaphysical (Kail, this volume). Like his 'educator' Schopenhauer (but without the metaphysics of a world will behind a merely apparent world of *Vorstellungen*), he regards the human as continuous with animal nature, thereby bridging or exploiting the intersecting set of the language games of animal and human. He also thinks that humans have cognitive capacities that most animals do not have. He insists that the human body is in need of critical attention (e.g. BGE 200), which implies that he assumes the possibility of change. In this sense then, such a *shaping* of what one is is possible, and what role consciousness and deliberation could possibly play in critical attention and self-control remains a much-debated question. When a self-system controls its drives, or when it is in an undecided state induced by competing, incompatible drives, he at times seems to assume that our intellect 'takes sides' (D 109).⁶ But is this 'taking sides' doing any work that contributes to the process and final outcome of what we would call a deliberation?

One thing seems clear: Nietzsche's naturalism is not first and foremost theoretical or metaphysical. As Han-Pile has recently shown, he actually intertwines naturalistic and transcendental aspects in his notion of 'naturalised a priori conditions' (2009). He is a practical philosopher-psychologist who wants to understand what is required for 'genuine proficiency and finesse in waging war with oneself (which is to say: the ability to control and outwit oneself)' (BGE 200).⁷

Even if free will in the traditional sense might have become an untenable position, and conscious, reflective judgement might not be as immediately efficacious as hitherto assumed, the possibility of critical attention, change, and some conscious control, matches our intuitions. Nietzsche's own philosophical efforts, too, seem premised on his belief that changing some of our beliefs, how we think and feel about the world, might make a difference. But in what sense and in what circumstances?

We find a first clue in the psychological explanation for the idea of responsibility and desert's free will (i.e. the kind of free will that allows us to attribute moral blame

⁶ 'to become aware that we are suffering from the vehemence of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in process in which the intellect has to take sides' (D 109).

⁷ This is just one example that shows that Nietzsche not only attacks a certain Cartesian substantive model of the self, which advocates a new understanding of the mind as physical, but also criticizes the mind as containing problematic tendencies, incorporated errors that in turn need our conscious attention. As Katsafanas (2005) has argued, Nietzsche thinks that becoming conscious of certain tacitly held beliefs, or changing certain beliefs, or adopting certain consciously held beliefs, will have a real effect. I discuss the latter in my Dries (2013).

and praise).⁸ Assuming sole responsibility for one's actions is really the attempt of a weak agent—who finds himself in a world he did not choose, often hostile to his desires and aims, propelled towards a future that is largely unknown and unpredictable—to transform himself into someone with a higher perceived self-efficacy, i.e. someone who has changed his belief about his capability to produce effects. A false belief nevertheless might have had beneficial effects in giving weak agents a different sense of agency, an authorship over their actions, and a feeling of independence.

With this in mind, I wish to turn to what I believe is Nietzsche's less well-known analysis of the phenomenology of freedom, an analysis that shapes his entire later philosophy.

3 Freedom As Function of Resistance (Resistance Axiom I)

Why do we, despite evidence to the contrary coming, for example, from the neurosciences, continue to believe in and use the term 'freedom'?⁹ Why is it that we so often *feel* free and feel it when we are not? Nietzsche puts forward an argument for our recalcitrance in giving up this talk of freedom. One of his earliest notes on the subject gives away the direction his later philosophy will take. He writes: 'The pleasure in power is to be explained from the displeasure experienced a hundred times from dependency' (NL 1876–7, KSA 8, 23[63]).

It is in an 1885 notebook passage that Nietzsche provides a fuller description of the phenomenology of freedom and offers an explanation for this obstinacy of our belief in it. It is interesting and carries weight because Nietzsche went back to revise it. In its earliest version, this passage read as follows:

When we encounter a resistance and have to give in, we feel unfree, when we do not give in, free. It is this feeling of our more of force, which we name 'freedom of the will': our force, which compels, against a force that is compelled. (Notebook N VII 1, 1st version of NL 1885, KSA 11, 34[250]; the underlinings are Nietzsche's)¹⁰

⁸ The *causa sui* is the best self-contradiction that has yet been conceived, it is a sort of logical violation and unnaturalness; [...] The desire for "freedom of will" in the superlative, metaphysical sense, such as still holds sway, unfortunately, in the minds of the half-educated, the desire to bear the entire and ultimate responsibility for one's actions oneself, and to absolve God, the world, ancestors, chance, and society therefrom, involves nothing less than to be precisely this *causa sui*, and, with more than Munchausen daring, to pull oneself up into existence by the hair, out of the slough of nothingness' (BGE 21).

⁹ Sections 7.3 and 7.4 draw on my 'The Feeling of Doing: Nietzsche on Agent Causation' (Dries 2013).

¹⁰ I quote here the transcription of Notebook N VII I according to division IX.1 of Nietzsche's *Werke Kritische Gesamtausgabe* (KGW), ed. M.-L. Haase and M. Kohlenbach (Berlin, New York: Walter de Gruyter): 'Wo wir einem Widerstande begegnen und ihm nachgeben müssen, fühlen wir uns unfrei, wo wir ihm nicht nachgeben, frei. Es ist das Gefühl unseres Mehr von Kraft, welches wir mit "Freiheit des Willen" bezeichnen: unsere Kraft, welche zwingt, gegen eine Kraft, welche gezwungen wird.'

In this passage Nietzsche is silent on whether ‘feeling unfree or free’ and the ‘feeling of our more of force’ are conscious first-person mental states or not. When he revises the passage, he explicitly addresses this question and adds that ‘freedom of the will’ refers to the *conscious awareness* of effort or, better put, efficacious effort or awareness of self-efficacy within a resistance relationship. With his revisions, the passage reads:

When we encounter a resistance and have to give in, we feel unfree, when we do not give in *but compel it to give in to us*, free. *I.e.*, it is this feeling of our more of force, which we name ‘freedom of the will’: *the conscious awareness of our force compelling, in relation to a force that is compelled.* (Notebook N VII 1, 2nd version of 34[250]; Nietzsche’s underlinings, changes and additions in italics)¹¹

Nietzsche’s analysis of the phenomenology of doing suggests that our deeply embedded belief in ourselves as free agents originates in resistance relationships that a self-system is consciously aware of—from the immediate awareness of self-efficacy registered simultaneously with a resistance relationship. This awareness or self-feeling, Nietzsche claims, provides first-person access to, and feedback on, the relational status of a self-system.¹²

¹¹ ‘Wo wir einem Widerstande begegnen und ihm nachgeben müssen, fühlen wir uns unfrei, wo wir ihm nicht nachgeben *sondern ihn zwingen, uns nachzugeben, frei*. *D.h.* es ist das Gefühl unseres Mehr von Kraft, welches wir mit “Freiheit des Willen” bezeichnen: *das Bewusstsein davon, dass unsere Kraft zwingt, im Verhältnis zu einer Kraft, welche gezwungen wird.* I use the English ‘compel’ in my translation of the German ‘zwingen’, as the German describes a phenomenology that refers to psychological force, authority, and control rather than primarily any act of ‘forcing’ in a physical sense. As Reginster (2007) has shown with regard to will to power, the phenomenology of freedom that Nietzsche describes here depends on resistance and demands resistance of many forms, but not domination or *overpowering* (cf. also TI IX 38).

¹² Nietzsche had an acute sense of the importance of self-efficacy and perceived self-efficacy, i.e. someone’s conscious or unconscious belief about their capabilities to produce effects. As Bandura (1977, 1978, 1994, 2000) and others have since shown, self-efficacy levels are indeed very important and have been proven to ‘affect life choices, level of motivation, quality of functioning, resilience to adversity and vulnerability to stress and depression’ (1994: 81). They have further shown that the perception of self-efficacy can change through social persuasion, but depends at heart on ‘inferences from somatic and emotional states of personal strength and vulnerabilities’ and on ‘mastery experiences’ (81). Nietzsche assumes that the first-person awareness of strengths and vulnerabilities provided an adaptive advantage of sentient organisms over non-sentient ones and led to certain valuations that are expressed in our primary and learned drives. Hunger, the drive to eat, or thirst, the drive to drink, would thus have their origins in an awareness of weak self-efficacy that results from a lack of nutrition or water. The minimal self is for Nietzsche a kind of feedback control system that, put crudely, we can picture like a thermostat. A drive is conceived as a certain need that, if not satisfied, puts the sentient self-system in a negative state or a state of tension that motivates sense-making and agency. Once a drive-induced affective orientation leads to the satisfying of the drive, a kind of equilibrium or homeostasis is reached until the next need arises and puts the system back into disequilibrium (on empirical evidence for homeostasis and its significance for neurobiology and psychology, see e.g. Damasio 2010). While little could be done to affect or change primary drives, Nietzsche targets many secondary, or learned drives which should be subjected to genealogical investigation, and, if no longer approved of, e.g. because of their nihilistic, ascetic tendency, should be unlearned (see, e.g., D 109, where Nietzsche describes how one can unlearn or at least control drives). Based on first-person experience, Nietzsche forms the hypothesis regarding an ‘instinct of freedom’ (GM II 17 and 18) or, better, a drive for efficacy or efficacious agency that, in the same way thirst monitors hydration levels, monitors the self-system’s overall resistance relationships, i.e. its levels of overall efficacy

Nietzsche's writings show that he forms a kind of explanatory hypothesis about self-systems: our sense of agency and ownership expresses (we might even say with Damasio 'is tracking') a self-system's overall resistance relationships, i.e. its levels of efficacy as an embedded agent, its relative strength and weakness in any given situation. And while this sense of agency might lead to the false belief in some fictitious Cartesian pilot, a soul substance, as evidence for a 'doer behind our deeds', Nietzsche nevertheless sees this sense of agency not simply as epiphenomenal but as intrinsically motivating, functioning like a standing desire: self-systems (unless they are a pathological case) need to feel themselves as efficacious agents. For example, in the *Genealogy* he uses this hypothesis to explain the emergence of what Nietzsche calls pejoratively 'slave morality'. Group or society formation, Nietzsche argues, prevents members from discharging their drives in the way they used to. As a consequence, these agents suffer from a lack of external resistance: 'Lacking external enemies and obstacles, and forced into oppressive narrowness and conformity of custom' motivates and leads to an internalization of man: 'All instincts which are not discharged outwardly turn inwards' (GM II 16). The absence of resistance scenarios that, according to Nietzsche's hypothesis, sustain the sense of self-efficacious agency, motivated the 'desire to give form to *oneself* as a piece of difficult, *resisting*, suffering matter' (GM II 18; my emphasis). This is how those agents, by turning *themselves* into resistance scenarios, restored their sense of self-efficacy and agency. The feeling of unfreedom, being merely constrained by society without being able to do anything about it, and the need for self-efficacious agency are assumed to have the motivational force needed for one to engage in actions that restore the latter.

It is this experiential feedback mechanism that fixes the meaning of 'free'. (Nietzsche argues that those who suffered most from heteronomy and unfreedom later misappropriated it. They devised a language game in which those who lacked efficacy restored it by endowing themselves with 'free will', 'absolute freedom', and 'absolute responsibility'.) Based on his phenomenological analysis, Nietzsche posits what I wish to call his 'resistance axiom' (RA) of freedom:

The degree to which a self-system feels free (is aware of itself as efficacious agent and author of its actions) is a function of a self-system's efficacious effort in its resistance relationships. Awareness of inefficacy motivates self-systems to restore their sense of agency.

The self-system feels free, feels itself an efficacious agent, when it is engaged in resistance relationships with which it can cope. The self-system feels unfree, is aware of other-agency, when it is engaged in resistance relationships with which it cannot cope.

as an embedded agent and its relative strength and weakness. On a primary drive level, unconscious and conscious awareness of efficacy, this is Nietzsche's speculative hypothesis, motivates the other drives to coordinate in such a way as to regain a certain level of perceived self-efficacy, but he also believes that the feeling of freedom depends in important ways on the level of resistance with which a self knows it can cope.

I do not have the space to go into contemporary discussions on these issues. All I want to show here is that Nietzsche's findings seem to be very similar to the results of Bayne and Levy's recent article on 'The Feeling of Doing':

Although the experience of authorship is... a central component of the phenomenology of agency generally, it appears to be particularly vivid in experiences of effort. [...] If the experience of authorship ever takes the experience of agent causation, we suggest that it is in contexts in which the experience of effort is particularly vivid. (2006: 63)

What have we established so far? Nietzsche seems to think that (1) our sense of self as agents and author of our actions is a function of a self-system's ability to cope with resistance, (2) the more effort and resistance a self-system can master, the higher its degree of self-efficacy and perceived freedom, and (3) if a self-system becomes aware of other-efficacy, thus feeling unfree, it is likely to motivate action that restores self-efficacy.

4 The Tenacity of Feeling Free

Nietzsche assumes that awareness of self-efficacy is deeply embedded. His naturalistic understanding of self-systems as biological organisms seems already to confirm the resistance axiom on an organic level, as a kind of proprioceptive, self-regulatory awareness.¹³ 'In the entire organism', he writes, 'there is constantly the overcoming of innumerable resistances/inhibitions.' Something like a feeling of agency arises: 'because we live in a state of innumerable individual pleasurable incitations, this expresses itself in the feeling of well-being of the entire person' (NL 1875, KSA 8, 9 [1]).¹⁴ Nietzsche hypothesizes that resistance relationships and awareness thereof exist already on the organic level. He describes what cognitive science today refers to as proprioception and interoception, our unconscious and conscious awareness of the states of our internal organs and muscles, which he believes result in a proto-conscious, pre-reflective sense of agency:

the sense of well-being as the feeling of *power* triggered by little obstacles: because in the entire organism there is constantly the overcoming of innumerable resistances/inhibitions,—this

¹³ Shortly after his remark on freedom he writes on experience in general: 'die Ermöglichung der Erfahrung, dadurch daß das wirkliche Geschehen, sowohl auf Seiten der einwirkenden Kräfte, als auf Seiten unserer gestaltenden, ungeheuer vereinfacht wird: so dass es ähnliche u. gleiche Dinge zu geben scheint. Erkenntniß ist Fälschung des Vielartigen u. Unzählbaren zum Gleichen, Ähnlichen, Abzählbaren. Also ist Leben nur Vermöge eines solchen Fälschungs-Apparates möglich. Denken ist ein fälschendes Umgestalten, Fühlen ist ein fälschendes Umgestalten, Wollen ist ein fälschendes Umgestalten—: in dem Allen liegt die Kraft der Assimilation: welche voraussetzt einen Willen etwas uns Gleich zu machen' (N VII 1, 3). See also: 'Das Mittel ist: die Einführung vollständiger Fiktionen als Schemata, nach denen wir uns das geistige Geschehen einfacher denken als es ist. Erfahrung ist nur möglich mit Hilfe von Gedächtniß: Gedächtniß ist nur möglich vermöge einer Abkürzung eines geistigen Vorganges zum Zeichen' (N VII 1, 6:12–26). For the pervasive function of signs, see Abel (1998, 2004, and this volume).

¹⁴ 'denn wir leben in einem Zustande zahlloser einzelner lustvoller Reizungen, das Wohlgefühl des ganzen Menschen ist der Ausdruck davon'.

victorious feeling becomes conscious as the overall feeling [*Gesamtgefühl*] of gaiety, ‘freedom’. (NL 1886–7, KSA 12, 5[50]53)¹⁵

It comes as no surprise that sentient self-systems cannot really free themselves from the freedom assumption, if the ‘*Freiheitsgefühl*’ (NL 1886–7, KSA 12, 5[50]48) denotes its overall sense of well-being and thus tracks the proper functioning of an organism’s resistance relationships.¹⁶

But it is not only on the level of the functioning organism that the sentient self-system is engaged in and seeking resistance activities. Nietzsche sees continuity with the more complex, higher-order cognitive functions of self-systems. Also, in its sense-making ability to see something as something, viewed from an intentional stance, Nietzsche sees confirmation that it is always dealing with and even seeking resistances. Even higher cognitive activities, such as imagining, naming, and thinking, he interprets as resistance activities that sustain our sense of agency. As long as a self-system is able to do so successfully, it feels ‘free’.¹⁷ This resistance awareness leads to what Nietzsche calls ‘our basic belief’: ‘[t]hat we are efficacious entities, forces’ (N VII 1, 6).¹⁸

This explains, Nietzsche thinks, why our belief in freedom is so recalcitrant, tenacious.¹⁹ Freedom, thus understood, has nothing to do with some mysterious, noumenal property that enables autonomous agents to start causal chains out of nothing. Nietzsche’s analysis of the phenomenology of freedom explains his belief that human beings—due to their unconscious and conscious awareness of how they are doing as natural, embedded, sentient self-systems—might be motivated and behave according to a drive or ‘instinct of freedom’ (as he writes in the *Genealogy*), and ‘function’ almost like feedback control systems: our basic sense of self as agent is

¹⁵ ‘das Wohlgefühl als das an leichten Widerständen sich auslösende Machtgefühl: denn im gesammten Organismus giebt es fortwährend Überwindung zahlloser Hemmungen,—dies Siegsgefühl kommt als Gesamtgefühl zum Bewußtsein, als Lustigkeit, “Freiheit”. See, for example, Damasio on empirical evidence for the hypothesis of a ‘protoself’ consisting of primordial feelings that ‘map’ and ‘reflect the current state of the body along varied dimensions, for example, along the scale that ranges from pleasure to pain’ (2010: 22, 272–3). For a recent defense of Nietzsche’s sensualism, see Riccardi (2013).

¹⁶ See also: ‘before “thinking” began, there must have been already a “composing”, the shaping faculty is more original than that of “thinking”’/‘bevor also “gedacht” wurde, muss schon gedichtet worden sein, der formende Sinn ist ursprünglicher als der “denkende”’ (NL 1885, KSA 11, 40[17]).

¹⁷ Neuroscience today distinguishes between a sense of agency (‘I initiated the action’) and a sense of ownership (‘It is my body’) (e.g. Gallagher 2000 and Tsakiris, Bosbach, and Gallagher 2007). Sense of agency and sense of ownership can come apart in involuntary action (e.g. a push) but are of particular importance for the investigation of pathological disorders like schizophrenia, Tourette’s, and alien hand syndrome (see also Zahavi 2005).

¹⁸ ‘Dass wir wirkende Wesen, Kräfte sind, ist unser Grundglaube’.

¹⁹ ‘Man hüte sich, von diesem ganzen Phänomen deshalb schon gering zu denken, weil es von vornherein hässlich und schmerzhaft ist. Im Grunde ist es ja dieselbe aktive Kraft, die in jenen Gewalt-Künstlern und Organisatoren grossartiger am Werke ist und Staaten baut, welche hier, innerlich, kleiner, kleinlicher, in der Richtung nach rückwärts, im “Labyrinth der Brust”, um mit Goethe zu reden, sich das schlechte Gewissen schafft und negative Ideale baut, eben jener *Instinkt der Freiheit* (in meiner Sprache geredet: der Wille zur Macht)’ (GM II 18).

the ‘monitoring’ of the self-system’s overall resistance relationships and leads to pro and contra affective and emotional states that structure a self-system’s orientation and action.

5 The Ascetic ‘Mechanism of Willing’

It is not only our immediate actions (drinking a sip of water right now because I’m thirsty and feeling rather weak) but also our larger cognitively held beliefs and world views that are related to our need for self-efficacy, Nietzsche argues. This is why he finds ‘morality’, in particular Christian morality, such an interesting case study. As has already been indicated, Nietzsche describes its emergence as a shift in beliefs motivated by the low sense of self-efficacy of the weak and it culminates in a world in which the ‘strong’ are ‘evil’ and the ‘weak’ are ‘good’ and are endowed with ‘free will’. While Nietzsche’s philosophy can be seen as a multifaceted attack on the consequences of this shift that instituted a nihilistic morality, it is also a shift that Nietzsche finds extremely interesting and from which he thinks one can learn a great deal, in particular about how agency and agential motivation work.

For Nietzsche, Christian morality is highly problematic. Low sense of agency and self-efficacy—their unfreedom—leads to an affective-cognitive state, namely resentment, which is expressive of their unsuccessful resistance against their oppressors. Such people maintain their sense of agency and self-efficacy only through the internalized struggle against their own drives and affects. Nevertheless, Nietzsche has great respect for the achievement. When a morality is so successful in disciplining self-systems, then we are bound to learn something by identifying its tools. And identifying its tools might be of particular importance if, as Nietzsche believed, the Christian moral framework has run its course, while it is still unclear what a non-ascetic agency might look like (as I said earlier, this is, I believe, Nietzsche’s central concern).

There is no room to rehearse the detailed arguments of GM and A here. Christian morality relied on values that are auto-destructive and lead to what Nietzsche calls ‘nihilism’, which denotes both a kind of disorientation and despair. While Nietzsche recognizes that the figurative ‘Death of God’ might make new agents a real possibility, he also anticipates an existential despair that arises when the highest values have devalued themselves, as well as ‘disorientation’ (I use Reginster’s terminology here). Values that were universally binding based on the authority of a deity enabled agents to orient themselves and motivates them to constrain their Pleistocene first-order drives. Real freedom lies not in reverting back to our older, pre-moral Pleistocene drive selves. Thus, the post-Christian might feel free, Nietzsche thinks, but he is no longer an agent and is in danger of merely becoming the playing field of his or her first-order drives. Nietzsche’s ideal of real freedom would also require a new,

non-ascetic self-disciplining and responsibility for one's actions. Thus it seems crucial to ask what we can learn from Christianity. How did the 'ascetics' manage to discipline themselves? Could the Christian mechanism perhaps be deployed in non-ascetic ways?

In a less well-known note, Nietzsche describes the mechanism of internalization that he thinks achieves the disciplining of drives through sublimation, with the help of the imagination:

The internalisation occurs when powerful drives, that are denied external discharge with the installation of peace and society, seek internal compensation, *with the assistance of the imagination* [my emphasis]. The need for enmity, cruelty, revenge, violence turns backward, 'withdraws'; in the pursuit of knowledge is avarice and conquest; in the artist the withdrawn power of the imagination and lying appears; the drives are transformed into demons, against which there is strife, etc. (NL 1887, KSA 12, 8[4])²⁰

The sublimation described is not specifically Christian. The pursuit of knowledge and art are nothing that Nietzsche disapproves of per se. Great artists are exemplars of drive discipline. Those who suddenly found themselves members of groups or societies that denied the discharge of their first-order aggressive and sexual drives sublimate the latter in their pursuit of knowledge and art. They used their imagination to simulate scenarios that would help them control the drives that would otherwise interfere with and endanger the groups or societies, by associating these drives with dangerous 'demons', which they feared, and against which they would be intrinsically motivated to struggle. Such images would generate con-affects that would provide the motivation to act to suppress and stifle their drives rather than satisfy them. They became early artists instead, and scientists, expressing their drives in ways that would no longer be harmful to their communities.

Already for this local society-enabling mechanism of sublimation, the imagination was extremely important in at least two ways. The imagination is used not only to generate powerful con-affects by turning drives into demons, but is also used to project and simulate specific, hard-to-achieve goals for the scientist or the artist, projects that would require immense effort. For their achievement, innumerable internal and external resistances would have to be overcome.

The religious and specifically Christian mechanism functioned by adding an additional imaginary layer around the local sublimations of artists and scientists. The imagination is used in order to simulate or project powerful scenarios such as

²⁰ 'Die V<erinnerlichung> entsteht <dadurch>, dass mächtige Triebe, denen mit Einrichtung des Friedens und der Gesellschaft die Entladung nach außen versagt wird, sich nach innen schadlos zu halten suchen, im Bunde mit der Imagination. Das Bedürfnis nach Feindschaft, Grausamkeit, Rache, Gewalt-samkeit wendet sich zurück, "tritt zurück"; im Erkennen-wollen ist Habsucht und Erobern; im Künstler tritt die zurückgetretene Vorstellungs- und Lügenkraft auf; die Triebe werden zu Dämonen umgeschaffen, mit denen es Kampf gibt usw.'

'eternal damnation' as punishment for the violation of absolute laws issued by an omnipotent deity. These go beyond any images hitherto available to the imagination and produce even more powerful affects in those who imagine such scenarios. Yet it was not only fear of eternal punishment but also the anticipation and imagining of eternal bliss and happiness that provided the affective framework to motivate a disciplining of the drives. The affects such produced enabled a drive ordering of almost all types: those who had already sublimated their drives through knowledge and artistry (the outcomes of the local mechanism) would order them even further according to Christian values (a global mechanism). And those who had been unable to sublimate their drives using the local mechanism would be more likely to abide by the laws and values of the religious and secular communities. In essence, what I am trying to show here is that the agents I have just described were able to achieve self-control through a mechanism I would like to call 'mental simulation', which allowed them to experience affects that would motivate them to control themselves.

I think one problem for Nietzsche certainly lies in the anti-individualistic, group-preserving aspects of the religious mechanism we have just described, which prevented any 'higher types' from flourishing. However, the *real* and indeed more pressing problem is actually the breakdown of the religious mechanism. In many respects, Christianity was a flourishing way of cultivating and constraining even those people who were not able to use the local mechanism to control and sublimate their drives and become artists and scientists, and would have lived only by their non-sublimated aggressive and sexual primary drives. 'Eternal damnation' and 'paradise' (or rather its loss) were able to do what society and its punishments were not: control or aggregate even those self-systems.

Since its decline, however, the modern self-system is in danger of a disaggregation. It carries too many contradictory lineages and drives to be unified (BGE 200). All those many drives, affects, and values that we implanted and were implanted in us are no longer constrained by the drive to please some higher power and the fear of eternal punishment, and modern self-systems have lost the framework that gave them their 'will', becoming unable to aggregate and discipline their many, and often opposed, drives. The breakdown of the mechanism of willing manifests itself in scenarios where A reflectively judges she ought to do ϕ and yet ends up θ -ing. The imagination is no longer providing the images, no longer awakens the affects that motivate constraint. She ϕ s, θ s, and σ s but without any overarching goal. Also, because former values and duties were based on absolute authority, all secular goals that come to mind, or are proposed by others, *feel* arbitrary and hardly *affect* her.

This captures in part what Nietzsche describes as decadence, for resistance scenarios if not ascetic are decadent. As Nietzsche describes it in EH, 'the true sign of decadence, being seduced by what is harmful, not being able to find your advantage any more (*das eigentliche décadence-Abzeichen, das Gelockt-werden vom Schädlichen, das Seinen-Nutzen-nicht-mehr-finden-können*)' (EH IV 8). Long-term goals lack motivational force. The self-system falls into a state of low efficacy, i.e. unfreedom and dependence,

because attempts to discipline or unify the drives are no longer successful, resulting in what Nietzsche calls the modern human being of ‘profound mediocrity’, who feels ‘unfree’ (NL 1886–7, KSA 12, 7[3]).

Against the previous ascetic ideals and the mediocre post-religious random ebb and flow of drives, new goals that unify the drive selves would again require the disciplining of the drives, as BGE 188 suggests:

What seems to be essential, ‘in heaven and in earth,’ is that there be obedience in one direction for a long time. In the long term, this always brings and has brought about something that makes life worth living—for instance: virtue, art, music, dance, reason, intellect—something that transfigures, something refined, fantastic, divine.

Again, Nietzsche tacitly argues from the resistance axiom. All the activities that made life worth living required ‘obedience’, the ordering of all drives in view of a goal. But how is it to be achieved? He regards this ordering as vital, particularly in light of the breaking down of the ordering mechanism Christian morality had provided. How can we make good his idea of agents that ‘proudly’ compose themselves, ‘have [their] affects, [their] For and Against, voluntarily’ (BGE 284), and reach ‘proficiency and finesse in waging war with oneself (which is to say: the ability to control and outwit oneself)’ (BGE 200)? Nietzsche is not very clear on how this is supposed to be possible. Rather than telling us how such new, non-ascetic agency might be possible, he describes ‘higher types’ that he regards as exemplars of a new ideal of agency. It is to these exemplars that I wish to turn to next.

6 Higher Types? (Resistance Axiom II)

It seems as if Nietzsche applies the resistance axiom of freedom as a kind of litmus test. In TI IX 38 ‘*My Conception of Freedom*’, Nietzsche analyses the ‘highest type of free man’ and provides important clues for his later psychology of freedom. ‘How is freedom measured in individuals and peoples?’, he asks:

According to the resistance that must be overcome, according to the exertion required to remain on top. The highest type of free men should be sought where the highest resistance is constantly overcome: five steps from tyranny, close to the threshold of the danger of servitude. [...] freedom exactly in the sense in which I understand it: as something one has and does not have, something one wants, something one conquers. (TI IX 38)

For the highest types, resistance is not the exception but rather what they value, their norm. Highest-type freedom requires resistance in two directions. In one sense, one must be close to ruling but not overruling because overruling would end the resistance. Simultaneously, one must be close to enslaving oneself and resisting because subjugation would end the resistance. The ‘highest types’, Nietzsche points out, have a standing desire for freedom, freedom is what they ‘want’, and they seem to know freedom’s secret: because freedom is what they want, they engage in projects

that require maximal effort, set themselves maximal resistance goals, and, as a result, they achieve the highest possible sense of agency. Nietzsche seems to assume that the anticipation of the highest freedom or self-efficacious agency has motivational force, which enables the higher types to discipline and order their drives in view of long-term resistance goals.

Goethe is described as another important exemplar. He is cited as someone who was able to set himself many goals and who displayed a higher, non-ascetic totality of multiple drives (NL 1887, KSA 12, 9[179]), which was not the result of a reductive unification. He ‘disciplined himself to a wholeness, he created himself’, he was ‘strong enough for this freedom’, achieved a non-reductive ‘tolerance, but not out of weakness’ (TI IX 49).

The ‘highest type of free men’ seek resistances and non-ascetically allow as many drives as possible to be active at the same time, but within a hierarchy so that they do not cancel each other out, and work towards one goal or towards the achievement of a common, hard-to-achieve project. A lot more could be said here but we need to press on because there is a second aspect to the resistance axiom that Nietzsche introduces in TI IX 38.

He observes that resistance does not only create the feeling of freedom and sustain the sense of self-efficacious agency. Maximizing and copying with effort and resistance also creates a particular kind of affective relationship to the ends thereby achieved, namely an attitude of appreciation or valuing. As he puts it at the beginning of TI IX 38: ‘The value of something does not lie in that which one attains by it, but in what one pays for it—what it *costs* us.’

The claim is that ends achieved or reached by effort due to resistance, ends for which one has to expend a lot, will be ‘dear’, and because they are dear, they will be ‘valuable’. The resistance axiom is therefore extended as follows:

- (i) The degree of freedom, sense of agency, and authorship is proportional to the resistance overcome; and
- (ii) The degree of freedom is coextensive to a pro-affect of appreciation or valuing. It follows that the:
- (iii) Value or meaning of something (object, activity, event, goal, life) is proportional to the resistance overcome.

Ends that do not involve resistance and exertion do not, or only minimally, make the agent feel free and efficacious, and ends reached rather than achieved will not be as dear or valuable to the agent. Consequently, the highest types would not only be those whose existence yields a sense of agency; also, due to maximal exposure to resistance, their existence is also maximally valuable to them. When something is dear and valuable it is something that one cares about and to which one ascribes meaning, towards which one will have a powerful pro-affect, an affirmative ‘yes’. It is immediately clear that the higher types are an attractive species for Nietzsche, who is critical of agents who only minimally sustain their sense of agency by ascetically

denying life; who, once the breakdown of the Christian motivational mechanism takes effect, will no longer even be able to do the latter, let alone affirm life non-ascetically.

This is all well and good, but isn't higher type agency and affirmation only possible precisely for those who actually are born as highest types? Or is Nietzsche pointing out a 'mechanism of willing' that the higher types employ naturally but that might nevertheless enable many agents to 'reach proficiency and finesse in waging war with themselves' and perhaps enable them 'to control and outwit themselves' (BGE 200)?

7 The Mechanism of Willing: Simulating Freedom

In this last part, I want to investigate briefly if Nietzsche's insights might be useful for a motivational theory that would help us make sense of his ideal of non-nihilistic agents who can set themselves long-term goals and projects and achieve an affirmative attitude towards life.

Let us return to the motivational mechanism Nietzsche detected in his negative exemplars: the mechanism always involved a process of deliberation or weighing that I called a *mental simulation*, which was in part cognitive (involving reflective judgement and beliefs) and in part non-cognitive (involving drives, affects, feelings, and emotions).²¹ During this mental simulation, the first-person perspective, namely affects that have a certain qualitative feel, a leaning for or against, plays an important role. Nietzsche assumes that societies and religions always provided the materials for the mental simulations their 'subjects' could use to 'affect' their self-systems in such a way as to achieve the right kind of self-discipline. This is what Nietzsche means by the casual remark: 'How did the possibility of eternal damnation take effect?' (NL 1881, KSA 9, 11[203]).²² Now, let us see if this motivational mechanism can really yield what it promises.

As we know from experience, agency takes place within an optional field. If we stay within the drive heuristics, then we can say that some options are related directly to our own drives and concomitant affects. Other options might be related to

²¹ I propose (but cannot defend here) that Nietzsche as both philosopher and psychologist assumes something like an early version of a simulation theory. For an early overview of the debate between theory and simulation theory, see Davies and Stone (1995); for a recent comprehensive defence of mental simulation, mainly focused on mindreading others, see Goldman (2006). On reasoning and emotional cognition as affective response to simulating one's own future states, see for example Thagard (2006). On some of the problems that simulation theory faces in explaining intersubjective understanding, see Gallagher (2007).

²² 'Wie hat die Möglichkeit der ewigen Verdammniß gewirkt!' In NL 81, KSA 9, 11[203] Nietzsche discusses the idea of eternal return. I take this passage to be an indication that Nietzsche conceives of eternal return primarily as a thought experiment, a thought experiment that can have real, affective-motivational effects, in the same way as the idea of eternal punishment had real affective, motivational effects and enabled agents to control their drive selves. I discuss this in Dries (forthcoming).

reflectively held beliefs, or values and affects that express the aims of others or even of institutions, groups, or societies, rather than our own. The first set of options, let us call them *autonomic* options, function like intentions: we would be motivated to carry them out. The second set of options, *heteronomic* options, lack this intentional force: we would not be motivated to act in the same way as we are when an option is autonomic, i.e. directly related to our own drives and their affective orientations. So how is it that we are sometimes able to go against our immediate autonomic options and act upon and carry out heteronomic options, e.g. a task that ‘we really did not feel like doing’?

One possibility would be to assume that there was some drive that takes charge and turns a heteronomic option into an autonomic option. For example, I have promised a friend that I would meet for dinner (a heteronomic option). Come Saturday night I no longer feel like going (no pro-affect or rather a con-affect ‘shows up’ when I think about or simulate going for dinner). Also, I can immediately think of things I would rather do (some autonomic option, for which pro-affects show up immediately). If I do not go, I anticipate my friend will be disappointed (a con-affect shows up, I don’t want to go to dinner, but I also don’t ‘feel like’ disappointing my friend). This is when I begin to deliberate, to simulate possible courses of action and their implications. When I think about it for a little longer, when I look at the website of the restaurant, I realize that they serve excellent *côte de boeuf* there. It turns out that I am presently hungry, and as I particularly like *côte de boeuf*—together with wanting to avoid disappointing my friend (after all, I think, what would I be without my friends and I shudder)—I suddenly witness the emergence of a motivational pro-affect and alongside that the thought that ‘it might not be such a bad idea, it might even be the right thing to go after all’. And off I go.

What has happened here? Our simulative ‘reasoning’ and weighing resulted in first the emergence of motivational force and then an action, when the agent, consciously or unconsciously, ‘listened’ to his affects. The intellect seems to play an important part in this simulation. There is some kind of cognitive penetration. It is the intellect that in moments of indeterminacy ‘offers’ different kinds of options or drives aims to different drives. Let us look at this in a bit more detail.

We mentally simulate an action and experience how our affects respond. (Sometimes people flip a coin in order to decide what course of action to choose. But they flip the coin not in order to let chance decide, but in order to find out how they affectively respond to the coin’s ‘selecting’ the one or the other option. This will show them what they really want.) When a deliberation/simulation does not result in any pro-affect, we might either respond that we ‘won’t do it!’ or we could also try to widen the scope of the simulation. For example, we often simulate consequences of our actions and think of the implications of our not carrying out the action. Eventually, we might stumble across something that motivates us via a strong affect or we ‘build up’ motivational force cumulatively, i.e. a number of contributing affect-reasons emerge.

The example I chose has, of course, very little to do with the kind of higher, life-affirming agency that Nietzsche seems to require. It is far from using and disciplining drives in view of a resistance-rich, hard-to-attain, creative long-term goal to which I subject myself. Can we imagine a scenario that would be more fitting?

A sense of agency and anticipation of increased self-efficacy are, Nietzsche believes, intrinsically motivating. What if we deliberated possible actions according to their potential to either increase or decrease our self-efficacy? If we had a standing desire for self-efficacy, it might motivate us to carry out certain actions, motivate us to select resistance scenarios that require extreme effort, because such actions would yield a higher agency freedom and the ends thereby achieved would really matter to us. Would this kind of simulation perhaps motivate specific activities (such as ‘finishing writing a book’ rather than ‘watching a TV series’)? And would it also induce justificatory thoughts about that selection (‘finishing the manuscript is the right thing to do’)?

Nietzsche thinks he might have found a solution to the puzzling question: ‘what “affects”, effects and rules the hierarchy of a self-system’s drives?’ If the constitutive drive for self-efficacy (‘freedom’) became conscious (rather than disguised behind altruistic, ascetic acts), then it could actually be used to ‘evaluate’ the possible constellations and options and generate pro- or con-affects depending on the increase or decrease of our agency freedom. As we saw, the Christian too relied on his instinct for agency freedom and self-efficacy, but not consciously. The ‘Christian’ cognitively disapproves of self-efficacy, instead practicing and thinking he ought to practice self-effacement. All the while, Nietzsche thinks, he is relying on the resistance axiom in a clandestine manner. If, on the other hand, the motivational force that issues from the desire for self-efficacy became known—what then?

Take another example: a recovering alcoholic is very tempted to give in to his strong and still active desire to have another drink. We know that many alcoholics relapse but we also know that some succeed and avoid relapse. How might they achieve control over their desire? Could a conscious simulation of the effort involved and anticipated self-efficacy, despite his strong affective orientation towards a drink (the world he perceives seems to present him with an almost inescapable path towards that next drink), motivate him to continue to resist his drive? Satisfying his desire for a drink would satisfy only a first-order desire and yield only a very minimal experience of freedom triggered by overcoming the resistance to the drive’s satisfaction. He anticipates a much greater overall self-efficacy from resisting a drink.

Nietzsche thinks we often have affective reactions that indicate that we already, without consciously doing so, evaluate a situation according to its potential to increase our self-efficacy or sense of agency. Pride is one of the affects he singles out in this respect. Human ‘pride’ can give rise to and motivate a ‘desire’ to combat any strong drive that is felt as ‘enslaving’ the self. In *Daybreak*, Nietzsche refers to ‘pride’ as father of the idea of freedom: ‘the doctrine of freedom of will has human pride and feeling of power for its father and mother?’ (D 128). In which sense is pride

used here? He sees pride as both the indication of a motivation from an immediate evaluation of a situation as possibly being too difficult for me, but as expressing the motivation to overcome this resistance, such as when we say someone is 'too proud to fail'. But it is also often an affect that anticipates achievement and follows exertion of some kind. This is part of what Nietzsche might have in mind here.

Let us return to our alcoholic. Supported by his pride against enslavement by his desire, simulating what it would be like to resist his craving for a drink, ordering water instead, and projecting what would be involved in ultimately resisting and starving the desire for alcohol he feels enslaved by (the simulation can go this far), would allow him to feel the possibility of a decrease in dependence and an increase in agency freedom and efficacy. He might, for example, imagine himself losing his job or his partner, he might anticipate himself in a weak, dependent, inefficacious position unable to successfully cope with life's challenges or unable to strive for other goals. But he might also get a sense of what it would feel like to resist successfully. This kind of evaluation might motivate him to avoid relapse and to recover.

Have we humans lost our pride? Nietzsche is a complicated internalist when it comes to practical reason. He does not find evidence that reflective judgements and beliefs are *by themselves* motivating. Something else is required to transform beliefs about what one could do into something like intentions that are already motivated. While most commentators insist that Nietzsche rules out the possibility of 'willing', this is not exactly true. Rather, as I have tried to argue, Nietzsche thinks that willing is a mechanism that is hard to spot because we have interpreted the phenomenology of willing in such a way that it misunderstands the full 'mechanism'. As he writes in GS 127: 'willing is actually such a well-practised mechanism that it almost escapes the observing eye'. For there to be 'will', Nietzsche suggests a simulation of attraction or repulsion is necessary:

first, in order for willing to come about, a representation (*Vorstellung*) of pleasure or displeasure is needed. Secondly, that a violent stimulus is experienced as pleasure or pain is a matter of the *interpreting* intellect, which, to be sure, generally works without being conscious of it (*uns unbewusst*); and one and the same stimulus *can* be interpreted as pleasure or pain. Thirdly, only in intellectual beings do pleasure, pain, and will exist; the vast majority of organisms has nothing like it. (GS 127)

So, there is willing after all. And while it had long been misunderstood or misinterpreted as *causa sui*, as a property of souls, of inner spectators with causal powers, etc., it is actually a mechanism that, once it is understood, might no longer remain and function as undetected as before. Once this complex drive psychology is disentangled, it is no longer a surprise that the later Nietzsche begins to imagine 'free' agents who can utilize what they now know of themselves, and learn how to perfect, how to 'have their drives and affects voluntarily'. Nietzsche is not at all opposed to the 'well-practised mechanism'. He is against the heteronomic, ascetic utilization by those who

were not aware that they felt like *causa sui* agents only because they had (been) turned against themselves. What if a self-system learned to no longer simply trust her affects, but instead learned to utilize them?²³

Understanding ourselves as complex, embedded self-systems who derive their sense of agency as efficacious agents from effort and resistance scenarios, from creative activities, might restore the mechanism of willing in the absence of the powerful illusions that have hitherto kept it going. Provided self-systems would learn how to value what is contingent and relative, with the same intensity they once loved gods, they might become agents who could set themselves circumspect goals, goals for which they are and feel responsible, out of a deep (affectively embedded) sense of concern, love, and responsibility for an unknown future as their task. 'Self-regulation is not achieved at once', Nietzsche writes in a note in 1881: 'The maximally free human being has the greatest feeling of power over himself, the greatest knowledge about himself, the greatest order in the necessary struggle among his powers' (NL 1881, KSA 9, 11[130]).

As Katsafanas (2009; 2013) has recently argued, Nietzsche proposes that agents are 'committed to valuing power merely in virtue of acting' (2009: 657), and that power should thus be seen as 'the constitutive aim of action'. It is for this reason that power enjoys a 'privileged normative status' (657) in Nietzsche's philosophy. While I agree with much of Katsafanas's compelling constitutivist reading, I hope to have shown that Nietzsche's interest in power, and its privileged normative status, cannot be fully appreciated unless it is viewed as part of his analysis of the phenomenon of freedom. Nietzsche values power, but most of all he values freedom, if correctly understood, and agents who experience freedom in their affirmative pursuit of the right kinds of ends. Suffice to say that none of this determines our substantive values and ends, or limits them to freedom and power for their own sake. What kinds of values we hold, or better, create, and what ends we deem worthy of pursuit will always have to be determined, negotiated, and then most likely renegotiated when circumstances change. Nevertheless, Nietzsche's analysis of freedom, as constitutive of and an important motivator for action, might help the kinds of things we are determine what kinds of good life we could be motivated to pursue.²⁴

²³ The drive monitors the self-system's activities along the lines Nietzsche suggests in a preparatory note: 'Does this stimulus increase or decrease our efficacy' (NL 1880-1, KSA 9, 10[F100]/9). The strong affects thereby produced provide the early components that give self-systems their 'will', albeit largely unconsciously. Also, a feeling of unfreedom is sufficient to feel oneself as an agent, albeit inefficacious. Recent research into the function of emotions suggests that emotions are mental states that express and thereby make available to a self-system very complex states that involve sense data, memory, tacit beliefs, etc. (e.g. Damasio 2003). They are therefore of vital importance as they can provide feedback and 'information' that would not otherwise be available to a self-system. On Nietzsche's anticipation of views in contemporary cognitive neuroscience, see Welshon (this volume).

²⁴ I am indebted to audiences in Geneva, London, and Oxford. The Cambridge Philosophical Research Colloquium devoted a session to an earlier version of this chapter. I am particularly grateful for comments by Raymond Geuss, Richard Raatzsch, Christian Skirke, Margaret Clare Churchill Ryan, and OUP's two anonymous reviewers.

Bibliography

- Abel, G. (1998), *Nietzsche: Die Dynamik der Willen zur Macht und die ewige Wiederkehr*, 2nd edn (Berlin; New York: Walter de Gruyter).
- Abel, G. (2004), *Zeichen der Wirklichkeit* (Frankfurt a. M.: Suhrkamp).
- Abel, G. (this volume), 'Consciousness, Language, and Nature: Nietzsche's Philosophy of Mind and Nature', in M. Dries and P. J. E. Kail (eds), *Nietzsche on Mind and Nature* (Oxford: Oxford University Press).
- Bandura, A. (1977), 'Self-Efficacy: Toward a Unifying Theory of Behavioral Change', *Psychological Review* 84/2: 191–215.
- Bandura, A. (1978), 'The Self System in Reciprocal Determinism', *American Psychologist* 33/4: 344–58.
- Bandura, A. (1994), 'Self-efficacy', in V. S. Ramachandran (ed.), *Encyclopedia of Human Behavior*, vol. 4 (New York: Academic Press).
- Bandura, A. (2000), 'Self-Efficacy: The Foundation of Agency', in W. J. Perrig and A. Grob (eds), *Control of Human Behavior, Mental Processes, and Consciousness* (Mahwah, NJ: Lawrence Erlbaum Associates).
- Bayne, T. J. and Levy, N. (2006), 'The Feeling of Doing: Deconstructing the Phenomenology of Agency', in N. Sebanz and W. Prinz (eds), *Disorders of Volition* (Cambridge, MA: MIT Press).
- Damasio, A. (2003), *Looking for Spinoza: Joy, Sorrow and the Feeling Brain* (Orlando: Harcourt Inc.).
- Damasio, A. (2010), *Self Comes to Mind: Constructing the Conscious Brain* (New York: Random House).
- Davies, M. and Stone, T. (1995), *Mental Simulation* (Oxford: Blackwell).
- Dries, M. (2013), 'The Feeling of Doing: Nietzsche on Agent Causation', *Nietzscheforschung* 20: 235–47.
- Dries M. (forthcoming), 'Decision-Making under Maximal Resistance', in D. Came (ed.), *Nietzsche on Morality and the Affirmation of Life* (Oxford: Oxford University Press).
- Gallagher, S. (2000), 'Philosophical Conceptions of the Self: Implications for Cognitive Science', *Trends in Cognitive Sciences*, 4/1: 14–21.
- Gallagher, S. (2007), 'Simulation Trouble', *Social Neuroscience* 2/3–4: 353–65.
- Gemes, K. (2013), 'Life's Perspectives', in K. Gemes and J. Richardson (eds), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press).
- Goldman, A. I. (2006), *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford: Oxford University Press).
- Han-Pile, B. (2009), 'Transcendental Aspects, Ontological Commitments and Naturalistic Elements in Nietzsche's Thought', *Inquiry* 52/2: 179–214.
- Janaway, C. (2007), *Beyond Selflessness: Reading Nietzsche's Genealogy* (Oxford: Oxford University Press).
- Janaway, C. (2012), 'Nietzsche On Morality, Drives and Human Greatness', in S. Robertson and C. Janaway (eds), *Nietzsche, Naturalism and Normativity* (Oxford: Oxford University Press).
- Kail, P. (this volume), 'Nietzsche's Explanatory Naturalism', in M. Dries and P. J. E. Kail (eds), *Nietzsche on Mind and Nature* (Oxford: Oxford University Press).

- Katsafanas, P. (2005), 'Nietzsche's Theory of Mind: Consciousness and Conceptualization', *European Journal of Philosophy* 13/1: 1–31.
- Katsafanas, P. (2009), 'Deriving Ethics from Action: A Nietzschean Version of Constitutivism', *Philosophy and Phenomenological Research* 83/3: 620–60.
- Katsafanas, P. (2013), 'Nietzsche's Philosophical Psychology', in J. Richardson and K. Gemes (eds), *The Oxford Handbook of Nietzsche* (Oxford: Oxford University Press).
- Katsafanas, P. (this volume), 'Value, Affect, Drive', in M. Dries and P. J. E. Kail (eds), *Nietzsche on Mind and Nature* (Oxford: Oxford University Press).
- Leiter, B. (2008), 'Nietzsche's Naturalism Reconsidered', *Chicago Public Law and Legal Theory Working Papers* 235: 1–40.
- Leiter, B. (2007), 'Nietzsche's Theory of the Will', *Philosopher's Imprint* 7/7: 1–15.
- Nietzsche, F. (1988), *Kritische Studien-Ausgabe*, ed. G. Colli and M. Montinari, 15 vols (Munich: Walter de Gruyter and Deutscher Taschenbuch Verlag).
- Nietzsche, F. (2001–), *Kritische Gesamtausgabe Werke*, Section IX, *Der handschriftliche Nachlaß ab Frühjahr 1885 in differenzierter Transkription*, ed. M-L. Haase and Michael Kohlenbach (Berlin: Walter de Gruyter).
- Owen, D. and Ridley A. (2003), 'On Fate', *International Studies in Philosophy* 35/3: 63–78.
- Reginster, B. (2007), 'The Will to Power and the Ethics of Creativity', in B. Leiter and N. Sinhababu (eds), *Nietzsche and Morality* (Oxford: Oxford University Press).
- Riccardi, M. (2013), 'Nietzsche's Sensualism', *European Journal of Philosophy* 21/2: 219–257.
- Richardson, J. (2004), *Nietzsche's New Darwinism* (Oxford: Oxford University Press).
- Richardson, J. (2008), 'Nietzsche's Problem of the Past', in M. Dries (ed.), *Nietzsche on Time and History* (Berlin, New York: Walter de Gruyter).
- Thagard, P. (2006), *Hot Thought: Mechanisms and Applications of Emotional Cognition* (Cambridge, MA: MIT Press).
- Tsakiris, M., Schütz-Bosbach, S., and Gallagher, S. (2007), 'On Agency and Body-Ownership: Phenomenological and Neurocognitive Reflections', *Consciousness and Cognition* 16/3: 645–60.
- Welshon, R. (this volume), 'Nietzsche, Consciousness, and Dynamic Cognitive Neuroscience', in M. Dries and P. J. E. Kail (eds), *Nietzsche on Mind and Nature* (Oxford: Oxford University Press).
- Zahavi, D. (2005), *Subjectivity and Selfhood: Investigating the First-Person Perspective* (Cambridge, MA: MIT Press).