

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/77830>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Exemplars as Evaluative Ideals in Nietzsche's Philosophy of Value

by

Jonathan Mitchell

A thesis submitted for the degree of Doctor of Philosophy (PhD)

University of Warwick, Department of Philosophy

October 2015

Table of Contents

| | |
|---|------------|
| Acknowledgements and Author's Declaration | iii |
| Summary of Thesis | iv |
| Abbreviations | v |
| | |
| Introduction to Thesis: Exemplars, Value and Selfhood | |
| 1. Prescriptivism | 2 |
| 2. Responses to Prescriptivism in Nietzsche Scholarship | 10 |
| 3. Normativity and Exemplars | 33 |
| | |
| I The Critical Ideal: Free spirit as Exemplar | |
| 1. Nietzsche's Methodology in HA | 42 |
| 2. Nietzsche's Critique of Metaphysics | 52 |
| 3. Moral Responsibility and Moral Sensations | 88 |
| 4. Nietzsche's Accounts of Guilt | 103 |
| | |
| II The Ethical Ideal: Zarathustra as Exemplar | |
| | |
| 1. The Speeches of the Prologue | 115 |
| 2. Nietzschean Self-overcoming | 128 |
| 3. Further Objections to the Ethical Ideal of Self-overcoming | 159 |
| | |
| III The Meta-Ethical Ideal: Future Philosopher as Exemplar | |
| 1. Value-Creation as Radical Choice | 172 |

| | |
|---|---------|
| 2. The Appeal to the Content of Conscious-Experience | 179 |
| 3. Nietzsche's Theory of Value | 186 |
| 4. The 'Noble' Mode of Valuing | 221 |
| Conclusion to Thesis: Connections and Further Research | 233 |
| Bibliography | 244 |

Acknowledgements

I would like to thank a number of people for their help and support during the writing of this thesis. First, I would like to express my gratitude to my supervisor Professor Poellner (University of Warwick) for his initial enthusiasm about the project and for providing consistently insightful commentary of numerous drafts of the various chapters. Secondly I would like to thank the University of Warwick for their award of a full scholarship without which I would not have been able to commit myself as fully to this project as I did. Thirdly I would like to thank my mother, and my two brothers (Robert and David), for having read through various parts of this project and discussed many of the ideas on a number of occasions. Finally I would like to thank my partner Charlie for her consistent encouragement, without which I would not have completed this thesis.

Declaration

I declare that (a) none of this material has been published prior to submission, (b) that this thesis is entirely my own work and (c) that it has not been submitted for a degree at any other university.

Summary of Thesis

The aim of this thesis is to provide a systematic account of Nietzsche's philosophy of value by examining his exemplars. It will be argued that these exemplars represent his favoured evaluative practices and therefore illustrate what I will call his evaluative ideals. The thesis will be structured in three chapters, each examining a different exemplar that emerges from a particular period of Nietzsche's work. Proceeding in this way will allow me to examine what I take to be three strands of his philosophy of value; the critical ideal through the exemplar of the Free Spirit, the ethical ideal through Zarathustra, and the meta-ethical ideal through the exemplar of the Future Philosopher. These standpoints, it will be claimed, reflect Nietzsche's central insights about *what* we should value and the *way* in which we should value, and are in this sense his evaluative ideals. Moreover, in doing so I will also attempt to provide some key insights on Nietzsche's reasons for his evaluative preferences, as given through these exemplars as evaluative ideals.

Abbreviations of Nietzsche's works

A = *The Antichrist* (Cambridge, 2005)

BGE = *Beyond Good and Evil* (Cambridge, 2002)

CW = *The Case of Wagner* (Cambridge, 2005)

D = *Daybreak* (Cambridge, 1997)

EH = *Ecce Homo* (Cambridge, 2005)

GM = *On the Genealogy of Morality* (Cambridge, 1997)

GS = *The Gay Science* (Cambridge, 2001)

HA = *Human, All Too Human* (Stanford, 2001)

KSA = *Kritische Studienausgabe*

NCW = *Nietzsche Contra Wagner* (Cambridge, 2005)

TI = *Twilight of the Idols* (Cambridge, 2005)

UM = *Unfashionable Observations* (Stanford, 1995)

WLN = *Writings from the Late Notebooks* (Cambridge, 2003)

WTP = *The Will to Power* (Vintage, 1968)

Z = *Thus Spoke Zarathustra* (Cambridge, 2006)

Note on translations

Translations from the KSA are my own or else otherwise noted. When quoting from Nietzsche's published works I have used the translations as listed above.

Introduction to Thesis: Ideals, Values and Exemplars

The aim of this thesis is to provide an account of Nietzsche's philosophy of value by examining his exemplars.¹ However, these figures will not be presented as loosely defined characters that he merely suggests we emulate. Rather, it will be argued that they represent his favoured evaluative practices and therefore illustrate what I will call his evaluative ideals. The thesis will be structured in three parts, each examining a different exemplar that emerges from a particular period of Nietzsche's work. Proceeding in this way will allow me to examine three strands of his philosophy of value; the critical ideal through the exemplar of the Free Spirit, the ethical ideal through Zarathustra, and the meta-ethical ideal through the exemplar of the Future Philosopher.

In this Introduction it will be argued that the standard interpretations of what Nietzsche's ideal amounts to are neither exegetically nor philosophically satisfactory. As an alternative I want to suggest that we can understand the different strands of his ideal as expressed in certain evaluative standpoints, the content of which will then be the focus of my chapters. These standpoints, it will be claimed, reflect Nietzsche's central insights about *what* we should value and the *way* in which we should value, and are in this sense his evaluative ideals.

However, it could be questioned why looking at Nietzsche's exemplars is a suitable method for understanding his philosophy of value. Since one might argue that Nietzsche's views should be examined in abstraction from his idiosyncratic, in the case

¹ 'Philosophy of value' as used here should be understood to include (a) normative ethics; that is, a concern with first-order, substantive ethical questions (e.g., 'what should I do', 'how should I live', 'what values should I be committed to') and (b) meta-ethics; that is a concern with second-order, foundational questions about the status of ethical inquiry and values (e.g., 'are there objective values', 'can ethics have a rational foundation', 'what does it mean to be a person'). For this distinction between first-order and second-order questions in moral philosophy see Mackie 1977:15-17.

of exemplars ostensibly pedagogical, mode of presentation. Yet, looking at Nietzsche's exemplars is useful in virtue of the fact that he so often, although by no means always, presents his philosophy of value through exemplary figures (a claim that this thesis will hopefully be ample evidence for). That this methodology is pervasive should not be particularly surprising given his claim that 'it is not through knowledge but through practice and through [having] a model that we become ourselves' (KSA 9:7 [213]). So, it is reasonable to assume that if we were to overlook Nietzsche's exemplars, we might miss at least some of the detail he gives us concerning his evaluative ideals. Moreover, as I will try to explain in the later sections of this Introduction there is also a more specific rationale for the exemplar-methodology that concerns the normative scope of Nietzsche's project.

1. Prescriptivism

As stated above, the justification for reading Nietzsche's ideal as connected to the notion of evaluative ideals as represented by exemplars grows out of the problems with the central readings of his ideal. One way of framing these readings is as different responses to the problem of prescriptivism. In this section I will explain what is meant by prescriptivism and consider an aspect of Nietzsche's relation to it in terms of his ethical particularism. Section 2 will then examine the responses of the secondary literature.

Prescriptivism in ethics, as it will be defined here, holds that there are certain rules or standards that have to be followed if one is to count as a moral person, and that a particular conception of what it is to be moral has objective authority, which therefore applies to, and is binding on, all persons (although different prescriptivist standpoints have explained the grounds of this kind of authority in different ways, e.g.,

Christianity's 'Will of God', Plato's 'Form of the Good', Kant's dictates of what follows for rational agents from the application of 'pure practical reason').² In this sense prescriptivist ethical theories do not merely make a conditional claim, that given a person desires to be moral, they should follow certain standards or rules, which would take the form of hypothetical imperatives (i.e., 'if you want X, do Y'). Rather, they make the stronger, categorical claim, that everyone *ought* to follow morality regardless of whether they are motivated to or not, insofar as it is claimed there is a fact of the matter about what counts as being moral, namely what the prescriptivist theory says it is, and that morality is taken to be indisputably important.³

So, prescriptivist ethics present themselves as non-optional, and non-compliance with the relevantly prescribed standards is taken as a significant failing. To justify this categoricity prescriptivist ethical theories must make a strong claim to possess legitimate normative authority, which, by their own lights, both (a) justifies the prescription of those unconditional moral oughts, and (b) vindicates evaluative assessments of individuals in relation to those moral standards, making persons appropriate targets of moral praise and blame regardless of whether they are in fact aiming at being moral in the defined sense.⁴

² For a different, and more technical use of the term prescriptivism see Hare 1963.

³ Such categorical claims are therefore what J. L. Mackie called 'action-directing absolutely' (see Mackie 1977: 26).

⁴ This aspect of prescriptivism might be captured by Nietzsche's comment that such standpoints claim moral hegemony, declaring, "I am morality itself and nothing else is moral" (BGE 202). Yet, this not to claim that categoricity and universal applicability *per se* is Nietzsche's central objection to forms of morality he rejects. As Brian Leiter points out, the more central objection Nietzsche makes against 'morality in the pejorative sense' (the broad family of views that come under the scope of his 'critique of morality') is what he sees as its normative agenda of harming or restricting actual or potential human excellence, although obviously this will be tied to making prescriptions of moral conduct for all persons (see Leiter 2015: 91-110).

It is also worth noting a distinction in terms of the way in which the categorical oughts of prescriptivist ethics can be either *agent-neutral* or *agent-relative*.⁵ In the *agent-neutral* case we can say that what is being prescribed, whether it is a general moral standard or a specific action, does *not* make an essential reference to the agent to whom it is directed, such that the agent does not figure in the content of the prescription (e.g., ‘live as Jesus did’, ‘murder should never be committed’). Conversely in the *agent-relative* case, an essential reference is made to the agent to whom it is directed in the content of the prescription. So, for example, categorical imperatives such as ‘all persons ought to live according to their means’, or ‘everyone should do what affords them the most pleasure’, are of the form such that how exactly the prescription translates into action (i.e., how the prescription is carried out) will be non-trivially determined by facts about the agent. We can see this in the above two examples by noting that if people have different means, or attain pleasure by different ways, then such prescriptions will yield a diverse range of actions *relative* to the agent. However, to make clear that categoricity is not undermined in the *agent-relative* case we can remind ourselves that the *agent-relative* prescriptions are *not* hypothetical imperatives. Since a hypothetical imperative in *agent-relative* form would be, for example, ‘people who want to be happy should do what affords them the most pleasure’, which means that if the desire to be happy is, for whatever reason, just not part of a persons motivational set then they are not bound by this imperative, i.e., it is *not* categorical and universally applicable in the way that the *agent-relative* prescription ‘everyone should do what affords them the most pleasure’ clearly is.

⁵ Although originally couched in terms of ‘reasons’ rather than ‘prescriptions’ this distinction, and the idea of ‘essential or non-essential reference’ to the agent, is that of Parfit 1984, and is also found in Nagel 1998: 112-114.

This distinction (between *agent-neutral* and *agent-relative* prescriptions) might be relevant for Nietzsche's ethical ideal, in that, as Edward Harcourt notes, if all he 'wanted to reject was the very strong claim [the legitimacy of agent-neutral prescriptions], there would be plenty of room for him to prescribe ways of life to others', as long as they were what Harcourt calls status-adjusted prescriptions, i.e., agent-relative prescriptions.⁶ As we shall see though Nietzsche might be opposed to such agent-relative prescriptions also, since he can be read as sceptical about the kind of normative authority that would be needed to legitimate the categoricity and universal applicability of even *agent-relative* prescriptions like 'all persons *ought* to live according to their means' (I will return to questions about normative authority in Section 3 of this Introduction).

Finally, we can also note the possibility of a position we might call prescriptivism *without principles*, a kind of 'particularist prescriptivism', that is clearly different from the more standard types of prescriptivist ethics in which general moral principles, of some sort, are central (e.g., Kantianism, Utilitarianism (both Act and Rule), and various forms of Religious Morality, most obviously Christianity and Islam). Such 'particularist prescriptivism' might hold that whilst there are no general prescriptions or rules for ethical conduct applicable across a variety of contexts, that nonetheless in specific situations there are certain categorical oughts that have to be followed. In other words, such a position would demand the same conduct of all agents, regardless of other motivational considerations, given they are in the same specific situation, and in that sense is a type of prescriptivist ethics which tells people what they ought to do, even if what they ought to do in one instance cannot be abstracted into some generalized principle, applicable beyond the specific situation.⁷

⁶ Harcourt 2011: 270.

⁷ See Korsgaard 2009: 72-76 for some discussion of this idea and some of its problems.

The hope is that by including this caveat what I am calling prescriptivism captures a broad family of views that I want to suggest Nietzsche is opposed to. They all involve (1) a categorical dimension (2) universal applicability, and (3) rules or standards (either *agent-neutral* or *agent-relative*) that both tell you what it is you *ought* to do and allow for an assessment of your conduct in those terms, even if, as in the case of prescriptivism *without principles*, the relevant rule would not necessarily have to be generalizable into some principle like Kant's Categorical Imperative.⁸ An obvious example of a prescriptivist ethics that fits this definition, and one that is often a target for Nietzsche, is Christianity and its moral ideal. At least on a standard reading, the Christian ideal that is being aimed at is a redeemed self, where if a person repents for their sins in this life they are granted eudaimonia in the kingdom of heaven. Yet this ideal state is not promised merely on the basis of holding certain beliefs, for example that Jesus was the Son of God. Rather, what is required is that one lives the life that one ought to lead following certain moral rules and standards as prescribed by the Bible and the relevant Christian authorities (e.g., the Ten Commandments).⁹ Moreover, and befitting an ethical standpoint that claims to exhaust the moral domain, Christianity makes negative assessments, and warns of punishment, for those who fail to adhere to

⁸ There is at least one historically influential moral standpoint that is not obviously an instance of prescriptivist ethics as here defined, that of Aristotle (although see Mackie 1977: 39-41). However in reference to Nietzsche at least two considerations should be kept in mind that would, for him, rule out an Aristotelian position. (1) Aristotle's conception of persons which presents them as having a kind of metaphysical *essence* which determines, given the suitable upbringing, what their proper moral ends will be in terms of a kind of rational teleology. Nietzsche's anti-essentialism (or at least anti *metaphysical*-essentialism) is in evidence in a number of passages, the most pertinent being GM I 13, WTP 551, WTP 556. (2) Aristotle's virtuous person exemplifies a high degree of prudential reasoning in his search for 'moderate' mean states. For Nietzsche's critical views of the model of practical reason central to Aristotelian ethics see BGE 198, 262, WTP 60.

⁹ See Exodus 20:1-17 and Deuteronomy 5:4-21.

the relevant standards.¹⁰ It is with such considerations in mind that Nietzsche describes such moral religions as ‘ones that prescribe how we ought to live and gain a hearing for their demands with rewards and punishments’ (WLN 2 [197]). So, it is relatively easy to see Christianity as an instance of prescriptivist ethics, with its categorical, often agent-neutral, prescriptions of one code of conduct for all.

In contrast to prescriptivist ethics, Nietzsche often gestures towards a kind of ethical particularism.¹¹ For example he has Zarathustra say:

I never liked asking the way...I preferred to question and try the ways myself...That, however – is my taste: - not good, not bad, but *my* taste... “This – it turns out – is *my* way – where is yours?” – That is how I answered those who asked me “the way.” *The* way after all – it does not exist! (Z, ‘On the Spirit of Gravity’, 2, cf. WLN 2 [203])

In other passages Nietzsche expresses a similar position, although sometimes basing it on a sceptical claim about our knowledge of the *effects* of ethical prescriptions, questioning our ability to validate any ethical prescription because, he asserts, their results ‘are either invisible or indistinct’ (D 24). Presumably what Nietzsche means by this is that the effects just do not show up in our experience of what follows from an ethical action, or at least not in a clear enough way as separable from whatever else is present in that experience, such as to validate the prescription. So, for example, a

¹⁰ As quoted from Aquinas by Nietzsche in GM I 15, ‘the blessed in the heavenly kingdom will see the torment of the damned so that they may even more thoroughly enjoy their blessedness’ (Thomas Aquinas, *Summa Theologiae* Supplement to the Third Part, question XCVII).

¹¹ This is not to claim that Nietzsche’s ethics is particularist in the sense of the term that is associated with the work of Jonathan Dancy. Dancy defends the distinctive ‘particularist’ thesis, in debates about practical reason, that the same reason for a person both doing and approving of an action in a particular situation can also be a reason for that same agent disapproving of it and not doing it in a different situation (see Dancy 1993: 56). For a critique of this view see Raz 2002, ch.10.

consequentialist ethics that might locate the value of an action in its putatively good effects, and therefore try to construct a valid prescription on that basis, would, according to Nietzsche, be evidentially suspect insofar as those putatively good effects never figure in our experience in the way that they would need to (presumably as visible and distinct) in order to confirm or disconfirm the prescription. So, his position seems to be that ethical prescriptions based on consequentialist considerations can ‘neither be demonstrated nor refuted from their results’ (D 24).¹²

In a later passage Nietzsche sounds a similar position although moving away from considerations of effects to a point about the particularity of ethical actions, telling us that ‘every act ever performed was done in an altogether unique and unrepeatable way, and...this will be equally true of every future act...as one observes or recollects any action, it is and remains impenetrable’ (GS 335). On this basis Nietzsche concludes that ‘all prescriptions of action (even the most inward and subtle rules of all moralities so far) relate to only their rough exterior’ and therefore ‘can never be proven true by our actions because every act is unknowable’ (GS 335, cf. D 116).¹³ Put simply, ethical actions are, according to this view, just not the kind of things that can admit of non-distorting generalizations that could then be translated into concrete prescriptions, they are both too (a) context-specific and (b) cognitively impenetrable, to yield something of

¹² See also D 108. I am not sure whether Nietzsche is right here or not, since it might be responded that in at least certain cases we can infer with a reasonable degree of accuracy what the *relevant* effects of a prescribed action are. So, for example, someone can observe the effects of a particular instance of following the prescription ‘murder is wrong’, insofar as the person they wanted to murder is still alive. Although it could be argued that prescriptions that relate to affecting other-directed mental states (e.g. ‘you should do what makes people happy’) are ‘invisible’ or ‘indistinct’ in the way Nietzsche suggests. Perhaps what a consequentialist ethics owes us then is a thorough account of how it is able, with sufficient clarity, to observe and separate out the relevant ‘effects’ of a prescribed action from the complex web of events, including the mental states of oneself and others, that follow from any ethical action.

¹³ This idea also finds expression in Kant 1991: 80.

this sort. So combined with the ideas about effects above (D 24), Nietzsche's ideal seems to be a kind of ethical particularist, who, as Peter Poellner notes, realizes 'that the features potentially relevant for evaluative appraisal of an action are so manifold, fine-grained, and mutually modifying that no substantive generalizations, capturing all these features and yet also being equally applicable across different actual contexts, are to be had'.¹⁴ Whether this position is ultimately defensible is not my central concern here. Rather, with this particularist stance in mind, as contrasted with the prescriptivism it is opposed to, I want to draw out a tension that the former view generates with regard to the notion of a Nietzschean ideal.

First, whilst it is fairly clear that a prescriptive ethics that claimed to know 'the way' (Z, 'On the Spirit of Gravity', 2) would contravene a particularist standpoint, we might think that the reasons which motivate Nietzsche's opposition to traditional moral pedagogy also threaten to undercut the possibility of him, without blatant inconsistency, offering contentful normative ideals. Since the very suggestion of ideals to follow, even suitably Nietzschean ones, would, if they had any substantial content, threaten to betray the individuality of a particularist ethic.¹⁵ Second, a problem emerges concerning the resources we are left with for assessing even a particularist ethic if we grant Nietzsche part of his rationale for that stance, namely that ethical actions are 'impenetrable' and 'unknowable' (GS 335). Since, if true, it is difficult to see what we could appeal to in order to make a judgement as to whether an instance of a particularistic ethic had been realized or was being lived out. Put another way, we might wonder what, other than evaluative appraisals of actions, which aim at approximately correct descriptions, could

¹⁴ Poellner 2009: 153.

¹⁵ One response to this worry would be to say that a particularist ethic has no business providing contentful ideals for other people to follow anyway, which would be tantamount to giving up on normative ethics. That this cannot be Nietzsche's position seems clear from reading his texts, as they are replete with all kinds of ethical recommendations and normative judgements.

step in to provide the relevant evidence for highlighting an instance of this achievement. In this sense part of the rationale Nietzsche provides for this particularistic ethic threatens to make recognizing such an achievement in myself, and others, problematic. Perhaps a more promising route would be to moderate Nietzsche's claim, such that we might say that it is not that *all* ethical actions are entirely 'impenetrable' or 'unknowable' (although perhaps some are) but rather that we do not know, with the requisite degree of certainty and clarity, those features (things like intentions and effects) that are relevant in moving from an approximate evaluative appraisal of an action to a categorical prescription on the basis of it.

So, by highlighting prescriptivism, and its contrast with particularism, we can see how proposing contentful normative ideals might be problematic for Nietzsche. Both exactly what his ideals are offering, and just as importantly the way in which they are being offered, in contrast with traditional prescriptivist ethics, needs to be accounted for in more detail. It will be argued in Section 3 of this Introduction that this tension can be resolved, at least in part, by appealing to the role that exemplars play in representing Nietzsche's ideals. However, before this I will discuss the central responses found in the secondary literature.

2. Responses to Prescriptivism in Nietzsche Scholarship

2.1 Self-eliminativism

One response to the problems above would be to argue for self-eliminativism, rejecting the idea that achieving some Nietzschean ideal is a project anyone could consciously

aim towards.¹⁶ As such, contravening Nietzsche's particularist commitments would cease to be an issue in that there would be simply no suitably autonomous self to whom normative recommendations could be effectively directed. Brian Leiter argues for a version of this position, telling us that for Nietzsche:

There is...no "self" in "self-mastery": that is, no conscious "self" who contributes anything to the process. "Self-mastery" is merely an effect of the interplay of certain drives, drives over which the conscious self exercises no control...A "person" is an arena in which the struggle of drives (type-facts) is played out; how they play out determines what he believes, what he values, what he becomes. But, qua conscious self or "agent", the person takes no active part in the process. (Leiter 2015: 80)

On this reading, any experience of self-mastery, of control over one's actions or values, is epiphenomenal, the real causes being sub-personal drives from which first-personal conscious states are effectively dislocated. In this sense the self reduces to a bundle of drives, and anything we typically associate with the control of the conscious self, such as some effective role in determining beliefs or values, is a product of a merely fortuitous aggregation of this bundle or a particular drive dominating.¹⁷

However, Leiter is not just claiming that for Nietzsche the self reduces to a bundle of

¹⁶ Some version of this view was often associated with post-structuralist readings of Nietzsche. See Deleuze 1977, Derrida 1977, Foucault 2001. For a recent version of this view see Metzinger, 'no such things as selves exist in the world: Nobody ever was or had a self' (Metzinger 2004: 1).

¹⁷ Note though that Leiter is arguing for a kind of self-eliminativism, rather than a broader eliminativism about the mental, which would be a kind of physicalism. He says, 'I am inclined to think that Nietzsche does reject both eliminativism and physical reductionism about the psychological, but neither is at stake for the kind of skeptical view of the self that Nietzsche holds, which requires the autonomy of the psychological as an explanatory realm (see Leiter 2015: 9).

drives, but also arguing for a specific reading of those drives as immutable, sub-personal psycho-physical entities (what he calls type-facts) whose relation to each other is also established (typically pre-established) at the sub-personal level. This reading therefore blocks a position which could agree that the Nietzschean self is minimally a bundle of drives, but claim that a normative ideal could be attained by conscious direction or control over these drives, perhaps altering or changing a particular drive, or even the overall structure of the drives. According to Leiter's reading, such an activity is just not within the remit of conscious direction given the nature of said drives.

This reading, and its response to ideas about 'cultivating one's drives', seems well textually supported given that on a number of occasions Nietzsche is resolutely sceptical about our knowledge of these drives. In one such passage he tells us that:

However far a man may go in self-knowledge, nothing however can be more incomplete than his image of the totality of drives which constitute his being. He can scarcely name even the cruder ones...above all the laws of their nutriment remain *wholly unknown* to him. (D 119, my emphasis, cf. D 109)

So Nietzsche is not merely claiming that we are unaware of some drives, say those on the periphery of our psychology. Rather, he is asserting that it is the 'totality of drives', that which constitutes what we are and therefore presumably includes our most dominant drives and their organized structure, of which we are ignorant. Moreover, what Nietzsche calls their 'laws of nutriment', ostensibly what would have to be known in order to cultivate one's drives, is, we are told, 'wholly unknown' to the conscious self. Given this view on our epistemological relation to the drives, which I will call the *conscious-impenetrability thesis*, it is not surprising to find Nietzsche stating elsewhere that the ability 'to combat the vehemence of a drive...does not stand within our own

power; nor does the choice of any particular method; nor does the success or failure of this method' (D 109).¹⁸

It is worth noting though, that if this is Nietzsche's position on our epistemological relation to the drives (and so to with regard to Leiter's type-facts) then it might be problematic if, as Leiter claims, these unconscious drives are held to be not merely causally primary with regard to, say, a value or belief we hold, but also explanatorily primary.¹⁹ Since given the *conscious-impenetrability thesis* is it unclear what kind of explanations such unknown causal mechanisms in terms of unconscious type-facts really provide. The thought is that if we want to appeal to some psychological-token as explanatorily primary we should be able to give, in at least some central cases, a substantial characterization of the nature of what is occupying this role and how it relates to what we do experience. However, in doing so the *conscious-impenetrability thesis* would be impinged upon. Therefore it might seem Leiter (and Nietzsche also on this reading) can only hold onto one of the claims here; either unconscious drives are explanatorily primary, and so must be able to contentfully figure in psychological explanations, or they are consciously impenetrable, not both. Perhaps we could distinguish between two views here. 1. The stronger version of the conscious-

¹⁸ For passages that gesture in a different direction on the drives see D 382, D 560. Ultimately I am not convinced, given the conflicting textual evidence (within the same text no less), that Nietzsche has a consistent position on what the status of drives and our relation to them are when they are intended to have this sub-personal reference. It is telling that discussion of drives in these terms largely drops out of the published corpus after *Daybreak*, perhaps reflecting Nietzsche's own dissatisfaction with his conflicting ideas. Although Nietzsche continues to use the term 'drive' in a less technical sense, as referring to a recurrent, habitual, or particularly strong, desire that a person or group of people possess. Yet such drives are often specified by him in terms of an epistemically accessible object (usually an end of some sort) typically revealed at the personal level, such as the drive to *knowledge* or the drive to *punish* (see BGE 13, BGE 189, GM III 9). For an account of Nietzschean drives that attempts to overcome some of these difficulties see Katsafanas 2013b.

¹⁹ See Leiter 2015: 72.

impenetrability thesis I have been discussing, and 2. A weaker thesis which might claim that certain traditional exercises in self-knowledge, be they reflection or introspection, for whatever reason distort or do not always give us reliable knowledge with regard to the causal mechanisms and explanations of certain parts of our psychological life. However, such a claim might just as well be framed in terms of a distortion of our pre-reflective, more immediate, conscious engagement with the world rather than some, ultimately more mysterious, sub-personal psychology.

In sum, if we accept Leiter's view then the topic of normative ideals in Nietzsche, and the problems we saw it generate, becomes largely moot. Below I will consider some reasons why we should resist this reading. However, Leiter's view, and its textual basis, does seem to impugn the chances of articulating an exegetically satisfactory account of Nietzsche's ideal as involving some consciously directed process of altering drives, if they are supposed to have, at least primarily, a sub-personal status.²⁰

2.2 Nietzsche's ideal as an achievement

A central problem with readings like Leiter's is that they overlook a significant number of passages in which Nietzsche suggests that consciously held attitudes, judgments, and commitments, usually in relation to matters of value, non-trivially and at least some of the time positively, affect the status of the self in question.²¹

²⁰ As a response to the question of why we should think of Nietzschean drives as sub-personal, the response would be a) Nietzsche often frames them in this way, and b) if drives are understood to be primarily personal-level phenomena then they become indistinguishable from things such as recurrent conscious desires. Again for a different view see Katsafanas 2013b.

²¹ Leiter's position seems to generate this oversight with regard to the more constructive questions surrounding the topic of value in Nietzsche insofar as from the onset he wants to separate Nietzsche's naturalistic project from his project of 'creating values' (see Leiter 2015: 9).

So, for example, in the early works Nietzsche tells us that ‘he who wants to become free has to become so through his own actions and that freedom falls into no one’s lap like a miraculous gift’ (UM IV 11), and that ‘many people live shyly and humbly in the presence of their ideal and would like to deny it: they are afraid of their higher self because when it speaks, its words are demanding’ (HH 9 624). Elsewhere he talks of an individual who posits ‘his own ideal’ and derives from it ‘his own law, joys and rights’, expressing a power ‘to create for ourselves our own new eyes and ever again new eyes that are ever more our own’ (GS 143), such that ‘to be a self’, in a more positive sense involves estimating ‘oneself according to one’s own measure and weight’ (GS 117). In the same work he goes on to describe ‘human beings who are new, unique, incomparable, who give themselves laws, who create themselves’, suggesting the ‘purification of our opinions and value judgements’ and the ‘creation of tables of what is good that are new and all our own’ (GS 335). Sounding similar themes in *Thus Spoke Zarathustra* Nietzsche claims that ‘the creative self created respect and disrespect for itself’ (Z, ‘On the Despisers of the Body’), and asks if one is able to ‘give yourself your own evil and good and hang your will above yourself like a law’ (Z, ‘On the Way of the Creator’). Finally, in later works Nietzsche recurrently highlights themes of self-legislation (BGE 210, 211) and suggests some ideal of a ‘sovereign’ self who possesses ‘eagerness for great responsibilities’ (BGE 213), and is distinguished by being able to hold himself to promises (GM II 2, 3).²²

Nonetheless, Nietzsche clearly does hold a sceptical view of the self in some sense, and it is important to clarify what conceptions he rejects. Nietzsche often criticizes a

²² For other passages which contain similar themes see UM IV 3, GS 99, 290, TI I 41, TI IX 38. Nietzsche’s thought also contains concepts such as *amor fati* (love of fate) and the eternal recurrence, in which not only is a certain evaluative attitude being described, but in a similar way to many of the passages above, a person’s status in relation to some ideal is assessed in terms of such evaluative attitudes.

conception of the self as a substantial entity that stands behind its actions. He tells us that, ‘there is no such substratum; there is no ‘being’ behind the deed, its effect and what becomes of it; ‘the doer’ is invented as an afterthought, - the doing is everything’ (GM I 13). Given this passage, and others like it,²³ Nietzsche is often read as critic of metaphysically realist theories of the self, represented, at least in his view, by notions such as the Platonic Soul, the Cartesian *res cogitans*, the Kantian Noumenal Self, and the Schopenhauerian Intelligible Character.²⁴ Nietzsche’s argument seems to turn on a disagreement about what is warranted with regard to the positing of a self if we describe the contents of our conscious experience accurately at the pre-reflective level. His claim being that metaphysically realist theories of the self that posit substantial selves, in certain cases located in metaphysical realms which necessarily transcend our experience, do so on the basis of an ‘afterthought’ (GM I 13), that is, through a type of reflection or introspection that abstracts from or distorts more immediate conscious experience. So, if we account for what is present in experience accurately we apparently find that there is no self which stands as a ‘being’ behind the deed’ (GM I 13).²⁵

I want to set aside whether an argument of this kind would be successful in combating the notion of a self as a substantial metaphysical entity. Rather, as a more limited exegetical point, what should be noted is that Nietzsche’s rejection of the notion of a substantial, or metaphysically realist, self does not lead him to reject the notion of an ideal self. This point is in evidence in those passages highlighted above, and is made

²³ See also WTP 491, 485.

²⁴ See BGE P, 12, 16, 17, and 54.

²⁵ For more extended discussion on this passage see Pippin 2006a. In the case of the Kantian self matters might be more complicated than GM I 13 suggests, since what is at issue for Kant is not merely some ‘substantial’ self behind our actions, but rather a more complex synthetic unity of consciousness without which, it is claimed, self-identification would not be possible. As Sebastian Gardner puts it, ‘what needs to be explained in the case of the I is not some common-or-garden existential belief, but a very strong, highly distinctive and highly complex reflexive unity, a capacity for constituting identity in difference’ (Gardner 2009: 6).

even clearer in a note where Nietzsche tells us that:

It is mythology to believe that we will find our authentic selves...Rather, making ourselves, shaping a form out of all the elements – that is the task! Always that of a sculptor, a productive man. It is not through knowledge but through practice and through [having] a model that we become ourselves... (KSA 9:7 [213])²⁶

Clearly, the first part of this note rejects a conception of the self as something which we might uncover, that is a kind of substantial entity which is already constituted, but somehow hidden. Yet, this scepticism does not lead Nietzsche to reject the notion of an ideal self as a normative achievement of some sort. So, the self-eliminativism in Nietzsche's thought can plausibly be read as local rather than global, insofar as it aims at undermining certain specific concepts of the self but need not lead to abandoning all notions of ideal selfhood. Rather, it is plausible to think that Nietzsche adopts this sceptical stance in order to then have the space to reconstruct a more positive conception. So, Lanier Anderson seems right when he remarks that:

The insufficiency of such [self-eliminativist] readings for Nietzsche's purposes is especially glaring when we turn to the autonomous self he idealizes, which exhibits a stronger, self-generated form of unity that far outstrips a mere 'bundle' of drives...The self's relation to itself and its attitudes towards itself ground the central normative judgements of Nietzsche's philosophy. (Anderson 2012: 232)²⁷

²⁶ See also UM III 1, 'your true nature lies, not concealed deep within you, but immeasurably high above you'.

²⁷ See also John Richardson; 'this vocabulary of drives isn't enough for the explanations Nietzsche wants to give us – nor the new values he wants to offer us' (Richardson 2009: 136-137).

Providing an interpretation of how best to construe Nietzsche's ideal self in these terms, as a normative achievement of some sort, along with the problems it generates, will be one of the central topics of the rest of this Introduction.

However, before proceeding I want to note a concern with any account that turns to considerations of Nietzsche's ideal self as a normative achievement. Anderson frames the problem I want to highlight with themes of self-creation in mind, telling us that 'the activity in question could not be *self*-creation unless one did it oneself, but that very self (namely, oneself) is the thing that is supposed to be created, and thus should first come into existence *only through* the process'.²⁸ In this sense there might seem to be a paradox in the notion of an ideal self as an achievement. Yet, this problem only arises due to an overly literal conception of 'self-creation', as if the 'created' Nietzschean self has some kind of realist ontological status that was previously lacking. Contrary to such impressions I think we should not take 'self-creation', or indeed striving towards some Nietzschean ideal, to be a metaphysical activity, in the sense of bringing into existence some entity that was previously not there. Rather, as identified in many of the passages in which Nietzsche frames his ideal self, we are primarily being presented with an *evaluative ideal*, and as such there need be no paradox or vicious circularity to the idea that one must, to a certain extent, already be oneself prior to engaging in some such project. As Anderson notes in Nietzsche's consideration of the case of Goethe, 'Nietzsche assumes that Goethe was already some kind of self before he 'disciplined himself to wholeness': indeed he was *himself*, in the sense sufficient for the self-disciplining activity to count as his own'.²⁹ So, any 'achieved self', or achieved evaluative ideal, would have to be psychologically continuous, to at least some degree and in some way, with the self that carried out the project to warrant the achievement

²⁸ Anderson 2012: 230.

²⁹ Ibid.

being identified as *one's own*.³⁰

Nonetheless, we might think that what is also needed to make sense of the notion of the ideal self as a normative achievement *of one's own* is some conception of a 'minimal' self that has the capacity (i.e., self-control) to strive towards some ideal. Clearly Leiter's interpretation, as we saw, wants to question the idea that Nietzsche's 'minimal' self is so constituted that any such achievement could be reasonably identified as my own. Yet such a reading takes on the burden of explaining away those passages in which Nietzsche seems to suggest, to his selected readership at least, an evaluative ideal to strive for. So, one response could be that insofar the project envisioned by Nietzsche is primarily an evaluative one, that is one which concerns what one values and the way one values, that the kind of control that the 'minimal' self would have to possess would just be the ability to engage in re-evaluative projects, something that we might argue Nietzsche seems to simply assume some of his readers will be capable of, as implied by him suggesting it in the first place. However, such a response will strike the self-eliminativist reading as question begging, since the capacity for consciously directed self-control (evaluative or not) is what is in question, and with some textual evidence too.

Yet interestingly, if we look to BGE 12, we find Nietzsche both rejecting substantial conceptions of the self whilst at the same time expressing a resistance to at least certain kinds of naturalistic self-eliminativism. He tells us that after having got rid of the 'the *atomism of the soul*', i.e., those substantial or metaphysically realist conceptions of the self, 'there is absolutely no need to give up "the soul" itself...as often happens with naturalists: given their clumsiness, they barely need to touch "the soul" to lose it"' (BGE 12). Now, it might be the case that the 'clumsy naturalists' Nietzsche is referring to here are in fact eliminativists about the realm of psychological explanation, i.e.,

³⁰ A similar point will be made in Chapter 2 about the 'self' in self-overcoming.

materialists or, in contemporary terms, physicalists who want to reduce psychological explanations to physical explanations (i.e., physiology or brain chemistry), rather than those like Leiter who hold a sceptical view of the self couched in a idiom of psychological explanations (i.e., drives, psycho-physical type facts). Nevertheless Nietzsche goes on in this passage to suggest that ‘the path lies open for new versions and sophistications of the soul hypothesis...concepts like the “mortal soul” and the “soul as subject-multiplicity” and the “soul as a society constructed out of drives and affects”’ (BGE 12). Whatever exactly *these* conceptions of the self amount to there is no necessity to thinking that they could not involve the capacity for some kind of evaluative self-control, indeed the reference to the ‘soul’ being ‘constructed’ rather seems to imply something of the sort.

Nevertheless, I cannot here provide a more detailed account of this ‘minimal’ self, and will hold off on a more detailed consideration of the meta-ethical issues that it generates until Chapter 3. For the moment I will proceed on the assumption, one at least implied by those passages above and BGE 12, that Nietzsche has *some* conception of the ‘minimal’ self beyond a mere bundle of drives or desires, and that, at least for some people, it might involve the ability for some degree of consciously directed evaluative self-control. Although the reader should bear in mind that the detailed argumentation for this form of self-control has not been provided. Still, whatever this ‘minimal’ self amounts to it is something more general than the evaluative ideals Nietzsche’s philosophy of value as a whole is directed towards, and as Anderson puts it ‘one must attain something further to become a self in the stronger normative sense’.³¹ So trying to understand Nietzsche’s stronger normative conception will concern me for the rest of this Introduction, returning us to the problems of particularism and prescriptivism.

³¹ Anderson 2012: 230.

2.3 Formal accounts of Nietzsche's ideal

An approach taken by some commentators is to distinguish between non-formal and formal ideals, claiming that Nietzsche's ideal only concerns the latter. In this way Nietzsche could avoid the charge of inconsistency with regard to his particularist commitments insofar as nothing is being prescribed in the relevant sense; his normative ideals would only provide broad aims whereas the individual determines the substantial content. As such these readings specify only the most general conditions of Nietzsche's ideal.³² Some examples of this nature are found in his praise of Goethe as someone who, 'disciplined himself to wholeness, he created himself' (TI IX 49) and his description of the 'great and rare art' of giving 'style' to one's character, which involves surveying 'all the strengths and weaknesses that their nature has to offer and then fit them into an artistic plan' (GS 290). Elsewhere he claims that 'unity in multiplicity' indicates greatness, even to the extent of assessing 'value and rank according to how much and how many things someone could carry and take upon himself' (BGE 212). Given this one could interpret Nietzsche as holding what Alexander Nehamas calls a 'formalist approach to character'.³³ As was suggested in Section 2.1 the prospects for construing this approach in terms of some activity relating to sub-personal drives is problematic. Here I want to consider approaches whose prospects seem better.

Ken Gemes claims that 'the idea of holding a multiplicity of aspects within a single coherent manifold is...the core to Nietzsche's account of what it is to be a person',³⁴ and Nehamas that 'our main object of concern with them [Nietzschean selves] becomes

³² See Gemes 2009: 33-50, Janaway 2009: 51-68, Nehamas 2001: 255-280, Hurka 2007: 9-31 (specifically Section 3).

³³ Nehamas 2001: 275.

³⁴ Gemes 2009: 43.

the overall manner of what they do, the very structure of their minds, and not primarily the contents of their actions'.³⁵ The claim is that what Nietzsche is most interested in is some kind of psychological wholeness as an achievement over and above any bundle of drives or desires, and that, as least some of the time, this is to be measured in terms of the quantity of different 'aspects' that can be held together in a coherent way. Given some of Nietzsche's own comments, it is reasonable to think that he might hold this view. However, as a satisfactory account of what his ideal amounts to, and on independent philosophical grounds, it faces a number of problems. Aside from being vague about exactly what those 'aspects' are, and how an individual might set about achieving such wholeness, we can think of a number of counter examples of types who seem to fulfill the formal criteria, but nonetheless, as Poellner notes, neither 'appeal to us...nor to Nietzsche'.³⁶

For example, the ascetic priest from the *On the Genealogy of Morality* seems to possess formal coherence, exhibiting what appears to be an established psychic hierarchy and even a certain kind of unity in diversity that allows him to carry out complex projects and be identified as a distinctive type. Nietzsche describes him as engaged in 'rule over the suffering...his instinct directs him towards it and his own special skill, mastery and brand of happiness are to be had in it...he has to be strong, more master of himself than of others, actually unscathed in his will to power' (GM III 15).³⁷ So, the ascetic priest is not obviously self-divided or a bundle of desires. Rather,

³⁵ Nehamas 2001: 277. Nehamas also claims that Nietzschean unity is modeled after the kind of unity we find in 'literary characters' (Nehamas 1985). For a critique of this view and why it is most likely not Nietzsche's see Leiter 1992: 275-290, and more recently Harcourt 2011: 256-284.

³⁶ Poellner 2009: 154.

³⁷ Hurka's additional claim that what distinguishes the formal achievement in question is whether it issues from 'strength and self-confidence' in contrast to 'weakness and resentment', like Janaway's that 'Nietzsche regards the presence of strong drives as characteristic of the

he appears to exhibit a kind of self-control which allows him to carry out a number of ostensibly difficult projects such as being central to the overturning of the aristocratic evaluative framework (GM I 6), and re-interpreting the suffering of humanity into a complex reflexive state that is said to avoid the latter giving up on life entirely; he exploits ‘the bad instincts of all sufferers for the purpose of self-discipline, self-surveillance and self-overcoming’ (GM III 16).³⁸ Yet, regardless of formal characteristics the ascetic priest is evidently not Nietzsche’s ideal. A similar criticism can be made of Thomas Hurka’s suggestion that ‘transformation of the world’ is the relevant formal criteria, such that for Nietzsche ‘the greatest individuals are those creators of new values who fix the course of life for millions of humans far into the future, and they are marked precisely by the extent of the goals they achieve’.³⁹ Since presumably St Paul, and the other early church leaders, who established a ‘new ruling and a church’ (WTP 167) could be said to fulfill this criteria and nevertheless are criticized by Nietzsche.⁴⁰ Criticisms that are based, one expects, on the content of their ‘new values’ (a point I will return to below), which whilst having been historically significant are clearly, for Nietzsche, problematic.

Perhaps the formal approaches could respond to such counter-examples by being more specific about the achievement in question. For instance, formal coherence as a kind of harmony or psychological consistency, which seems to be Nehamas’s view on what the relevant ideal is,⁴¹ is quite broad, and can be distinguished from those passages

greatest or healthiest type of human individual’, will not work in the case of Nietzsche’s ascetic priests, since whatever else an ‘unscathed’ will to power is, one assumes it involves strength, and ‘strong drives’, of some sort. See Hurka 2007: 26 and Janaway 2012: 188.

³⁸ For a subtle account of Nietzsche’s ascetic priests that emphasizes their important role in the ‘slave revolt in morality’ see Anderson 2011: 24-55.

³⁹ Hurka 2007: 24.

⁴⁰ See GS 139, A 51, WTP 171, 175, 205.

⁴¹ Nehamas 1985: 7. See also Robertson 2012: 106.

in which Nietzsche's ideal of 'unity in diversity' is a more complex state, which involves, as Harcourt puts it, 'tolerating *disharmony* – being such that conflicting tendencies can coexist in oneself *without* succumbing to the temptation either to disown one or more of them, or to harmonize them'.⁴² So, it would be an achievement of this kind that Nietzsche is primarily suggesting.⁴³ Perhaps then, insofar as the state being aimed at is more complex in character, we should not assume it would be exemplified by the ascetic priests (although more argument would have to be given as to why they do not fit this amended criteria, after all they hardly seem to exemplify a kind of characterological harmony either). Still, what exactly the ideal of unity in diversity as 'tolerating disharmony' amounts is problematic, since we might wonder how, other than by engaging in self-deception, sublimation, or simply rejecting one side of a conflict and resolving oneself in some sense, opposing standpoints can be held together in a unified way.

Harry Frankfurt's discussion of wholeheartedness suggests one way in which we could think of such an achievement. According to Frankfurt wholeheartedness does not amount to eliminating conflicting desires or tendencies from one's psychic economy, but rather involves identifying with one side of the conflict, such that the agent comes to disassociate his *self* from what are deemed to be the conflicting tendencies. Frankfurt tells us that 'the conflict between the *desires* is in this way transformed into a conflict between *one* of them and the *person* who has identified himself with its rival'.⁴⁴ Such a

⁴² Harcourt 2011: 278. See also Janaway 2012: 188-189. Although this might conflict with Nietzsche's praise of a kind 'Goethean' serenity as a kind of harmony of the soul (WTP 283).

⁴³ For passages which sound this view rather than the harmony view, see BGE 200, GM I 16, EH, 'Why I am So Clever', 9. Nietzsche's praise of self-overcoming and suggestion that some of his readers re-evaluate certain values is also problematic for 'harmony' readings. Presumably both of these projects would destabilize the 'coherent self' and potentially involve actions which, for a time at least, would result in forms of self-division and self-conflict.

⁴⁴ Frankfurt 1998: 172.

process would therefore involve, as David Velleman puts it, disassociating from ‘unwelcome desires’ in such a way as to expel ‘them from the self without necessarily eliminating them entirely’.⁴⁵ However, as Harcourt frames it, Nietzschean toleration of disharmony involves resisting the temptation to disown one side of the conflict. So whilst Frankfurt’s wholeheartedness might not involve extirpation of a particular desire from one’s psychic economy, it clearly involves a disassociation with certain desires which would amount to disowning them, and so seems like a narrowing of the boundaries of the self, whereas Nietzschean toleration of disharmony, or ‘unity in multiplicity’ (BGE 212), seems more like a broadening of those boundaries.

Nevertheless, we might ask how much sense the ideal of unity in conflict as something different from Frankfurt’s wholeheartedness makes. Consider for instance Christopher Janaway’s suggestion that Nietzsche’s ideal involves ‘a unity between elements that conflict...when the whole system is strong, rather than falling apart, the drives function together towards ends that are those of the individual as such’.⁴⁶ Putting aside reservations about the language of drives, presumably the idea cannot be that contradictory drives or strong desires, say the sex drive and the drive to abstinence, can function together towards the same end, since what end could they both pursue in conflict? Surely a drive or strong desire is individuated, at least primarily, by its end; the sex drives’ end being sexual gratification, and the end of drive to abstinence being the withholding of such gratification. The idea that an individual might have both these drives, or indeed any drives or strong desires that are antipodal, and nonetheless pursue ends that involve utilizing both, ‘functioning together’ as Janaway puts it, seems implausible or at least in need of further explanation.

⁴⁵ Velleman 2006: 342. Also see Velleman 2006: 330-360 for a critical response to Frankfurt’s idea of wholeheartedness.

⁴⁶ Janaway 2012: 188.

Moreover, one might wonder whether an individual who is, at least to some extent, willfully maintaining themselves in a state of psychological conflict, not ‘succumbing to the temptation’ to ‘harmonize’ oneself,⁴⁷ might not be able to act. In this way such a formal ideal might result in *akrasia*, where the reasons for action that certain strong desires provide cannot be followed on pain of resolving the conflict. Or alternatively if one desire was eventually to win out, we might wonder what contribution the person could be said to have made, and why they should identify with the desire that wins out rather than the one that loses. Remember that one of Nietzsche’s claims is that to be a self in a stronger sense involves living ‘according to one’s own measure and weight’ (GS 117), yet the kind of ambivalence involved in refusing taking a stand on one’s conflicting desires might seem to undermine this ideal (i.e., that I am living by *my own* ‘measure’ and ‘weight’ rather than that of some desire or drive), or at least make my relation to it ambiguous, as seen in Janaway’s reference to ‘a whole system’ as strong rather than a *person* as strong.⁴⁸

Moving onto a second line of criticism against formal approaches I want to suggest that regardless of how exactly the formal ideal is specified, or whether such accounts could meet the above objections, they do not do justice to central features of Nietzsche’s own assessments of persons in non-formal terms, a feature highlighted by the grounds of the mature Nietzsche’s evaluation of Wagner. Since his negative assessment is not based on Wagner becoming less unified, or not exhibiting unity-in-diversity when he previously did. Rather, the criticism that Nietzsche levels is that there is something about the content of what Wagner’s values, what we might call his first-order evaluative standpoint, which Nietzsche describes as expressing a ‘corrupted taste’ (CW 5), as exemplifying a ‘crisis in taste’, regardless of the fact that he was still ‘something *complete...whose every feature was a necessity*’ (CW 7). So it is something about what

⁴⁷ Harcourt 2011: 278.

⁴⁸ See Frankfurt 1998: 164 for discussion of this problem.

the content of Wagner's evaluative standpoint expresses that Nietzsche objects to, and not merely insofar as it is indicative of some formal defect.

More specifically it could be argued that what Nietzsche takes himself to be assessing here, and in other similar examples, is what we might call, following Poellner's remarks on the use of the term 'taste' in Nietzsche, a 'pattern of conscious affectivity'.⁴⁹ In these contexts then the target would be a particular kind of responsiveness to the world, a kind of affective-evaluative sensibility, an assessment of which will necessarily have to make reference to the *qualitative content* of the evaluative standpoint in question (i.e., what it is about and what it is aiming at). So in Wagner's case again, he highlights 'the convulsiveness of his affects, his over-charged sensibility, his taste that craves stronger and stronger spices, the instability that he disguises as a principle, and not least his choice of heroes and heroines' (CW 5).⁵⁰ This dimension to Nietzsche's evaluative assessments of persons, as concerned with both (a) critiquing, by his lights, objectionable first-order evaluative standpoints in terms of the kind of 'taste' they are expressive of, and (b) pointing towards or seeking to cultivate a 'higher', or ideal, kind of 'taste', whilst in need of much further elaboration, is overlooked by formal approaches.⁵¹

What the preceding discussion shows is that formal characteristics are not sufficient for the presence of Nietzsche's ideal. Nehamas concurs, remarking that 'consistency may not in itself be a condition sufficient for its presence'.⁵² So perhaps we could adopt the weaker position and hold that they are necessary for it. Yet, a number of additional

⁴⁹ See Poellner 2007: 231. For other examples of this mode of evaluation see, GS 39, GS 302, BGE 43, BGE 186, BGE 233, BGE 263, GM III 19, A 13, A 19.

⁵⁰ See also WTP 849, WLN 2[34], CW 8 and CW postscript; 'This is precisely what the case of Wagner proves: he won over the crowds, - he ruined taste, he ruined our taste even for opera'.

⁵¹ I will try to elaborate these ideas further at the end of this section but will do so more fully in Chapter 3, Section 4 of this thesis.

⁵² Nehamas 2001: 266-267.

worries arise with the formal approach with regard to its view of the relation between character on the one hand, and actions and values on the other, that might undermine even this more limited role.

Up to this point the formal view could be expressed as follows:

Formal Metric: A value or action is good for Nietzsche to the extent that it expresses or promotes the formal ideal. A value or action is bad for Nietzsche, if it inhibits or is not expressive of the formal ideal.

This position is close to the view of Hurka who tells us that Nietzsche ‘evaluates goals in terms of formal qualities...it is not its specific aim that determines an activity’s degree of worth, but how far that aim instantiates certain formal principles’.⁵³ However this metric for assessing first-order goals is unappealing for a number of reasons.

Firstly, we might think that a person cannot meaningfully aim at ideals such as unity-in-diversity or characterological coherence without already being engaged in a more concrete project, and because of this we might worry that the metric given above generates potentially irrelevant results for those who disagree about the value of the non-formal project in question. For example, given that someone values being a revolutionary they might aim for a certain consistency or even unity-in-diversity in that project. But if they did not already place value on the non-formal project then the aim of coherence or consistency would be irrelevant. Moreover, if we do not share a positive estimation of the particular project in question it is unlikely that the coherent, or otherwise ideal, nature of their engagement in that project will change our assessment of their worth. In fact we can imagine certain instances where, if we do not the project in question, inconsistency therein might be a redeeming characteristic. Therefore we might

⁵³ Hurka 2007: 23.

think that our assessments of the value of formal achievements are themselves conditioned by our stance on the relative worth of the first-order ends or projects being pursued.

Further, it might seem that the value of these substantive projects needs to be established first, and so has a kind of explanatory priority, before formal considerations can appear as relevant, and therefore inform appraisals of a person's character. In this sense it might not merely be, as Nehamas remarks, that Nietzschean character is not independent from 'the *quality* of the actions of which it constitutes the pattern',⁵⁴ but more strongly that an evaluation of the qualitative nature of those actions and ends sets the terms of our receptivity to any such formal achievements, i.e., whether or not they are achievements we should care about.

Secondly, whilst when assessing other persons we might be able to, by means of abstraction, overlook the content of particular actions and ends, for the individual themselves, as they are 'living through' any project, it seems implausible that they could aim towards such a formal metric as an ideal. Since doing so would undermine the sense of why it is they are pursuing any particular end rather than others that fulfil the same formal criteria.⁵⁵ As Harcourt notes, 'if it really doesn't matter to one that one has this or that set of long-term ends, it's doubtful that there is room for one to think of these ends as good, since to think of them as good in some way would surely be for them to start mattering'.⁵⁶ In this sense formal metrics of first-order goals and ends might create problems of motivation. For example, if a person really did not fundamentally care, in any non-instrumental sense, about the content of the ends they

⁵⁴ Nehamas 2001: 266-267.

⁵⁵ Presumably there will be a number of diverse substantive projects and ends that fulfil the formal criteria insofar as part of the point of formal ideals is that they allow for a wide-range of non-prescribed contents.

⁵⁶ Harcourt 2011: 271.

are pursuing, as Hurka's interpretation seems to imply,⁵⁷ that is if they were in no sense recognized as actual goods and therefore as in some sense intrinsically motivating, is hard to see how such a person would be drawn to any particular end that happened to fulfil the relevant formal criteria rather than any other that did likewise. The sense of pursuing any particular end would always be to some degree arbitrary (especially in reflection on those ends), and so an individual might, at any moment, be just as well motivated to pursue any other that satisfied the formal requirements.

Finally, it could be argued that treating first-order ends as either fundamentally or to some significant degree to be measured in terms of formal criteria is expressive of a certain kind of evaluative defect. Since, we might think that an evaluative perspective which makes a pre-judgement to *not* find any ends or projects intrinsically valuable (as worth pursuing in their own right), or indeed disvaluable (as not worth pursuing in their own right), is guilty of always treating first-order ethical goods as means rather than ends. As Aristotle writes, 'desire would be empty and vain...[unless] there is some end of the things we do which we desire for its own sake'.⁵⁸ So, this attitude might not only fit poorly with our experience of a wide range of evaluative activities (think of how impoverished a view of our reasons for aesthetic contemplation, or of our reasoning about evaluative commitments, the formal metric would give us) but also, as the Wagner example was intended to show, is not in keeping with aspects of Nietzsche's own evaluative practice which is responding to, and making assessments in terms of, non-formal features.

Given these worries it seems that formal approaches to Nietzsche's ideal are neither exegetically nor philosophically satisfactory.⁵⁹ Yet the former at least should not be

⁵⁷ Hurka 2007: 23.

⁵⁸ Aristotle 1998: 1094a 18-21.

⁵⁹ A different approach is taken by Bernard Reginster who makes a psychological version of will to power the standard by which actions and values are to be assessed, where Nietzsche's

particularly surprising given that Nietzsche's philosophy of value is regularly framed in terms of a 'rank ordering' of first-order ends and values as 'higher' or 'lower' in a way that often makes an essential reference to *qualitative content* of the evaluative standpoints in question.⁶⁰ It is in this spirit that the language of 'higher' and 'lower', and of 'good' and 'bad', is often framed by Nietzsche in terms of 'taste'. For example he describes his ideal as having 'refined senses and a refined taste; to be accustomed to the exquisite and most excellent things of the spirit' (GS 302), and it is an appeal to taste that putatively underscores certain evaluative judgements; 'what decides against Christianity now is our taste, not our reasons' (GS 132). Elsewhere he talks of having a 'different' and more 'subtle' taste as an *achievement*, expressed in 'a more delicate taste for joy, with a more tender tongue for all good things, with merrier senses' (GS P 4), and in later work that 'you need to have made considerable sacrifices for good taste; you need to have done many things, left many things undone for its sake...good taste needs to have provided you with a principle of selection' (TI VIII 47).⁶¹ Finally he has Zarathustra say that 'all life is disputing of taste and tasting! Taste: that is simultaneously weight and scale and weigher, and woe to all that would live without disputing weight and scale and weighers' (Z, 'On the Sublime Ones'). Understanding how exactly 'good taste' could provide someone with a 'principle of selection' amongst first-order ends and goals, and the conditions Zarathustra suggests for understanding

ideal is an individual who 'by virtue of engaging in the pursuit of power, is perpetually in search of new challenges to meet, of new overcomings' (Reginster 2006: 251). This reading is sufficiently distinctive to warrant dealing with separately, and since it also provides an interpretation of the notion of self-overcoming, which is the topic of the second chapter of this thesis, I will consider it there rather than in this Introduction.

⁶⁰ See BGE 213, 260, 268, 287, GM I 17.

⁶¹ Nietzsche also seems to think that there is a certain kind of evaluative truth to be had in virtue of possessing of a 'higher' taste, in evidence in where he equates 'a good, a refined sense of tact and taste...[with] reality' (A 59). For more detailed consideration of this theme see Chapter 3, Section 4.

any instance of particular ‘taste’, are complex questions that I will only be able to deal with properly later in this thesis.⁶²

Putting them aside for moment, I think we can nonetheless see that substantive considerations are essential to understanding Nietzsche’s ideal. Indeed it seems reasonable to think that having the right kind of ‘taste’, having a certain kind of affective-evaluative sensibility and by dint of this also certain evaluative commitments, is central to Nietzsche’s philosophy of value. Moreover, it is the importance of this qualitative dimension to Nietzsche’s ideals that makes prescriptivism, along with an elaboration of the reasons for these evaluative preferences, a genuine problem that needs to be addressed.⁶³ Formal accounts of Nietzsche’s ideal therefore avoid being prescriptive at the cost of potentially depriving Nietzsche’s philosophy of value of its substantive force, as Harcourt notes, they ‘certainly give us an interpretation of a non-prescriptive conception of how to live’ but are ‘too weak’.⁶⁴

⁶² The conditions for Zarathustra seem to be determining (a) what value has been accorded to a particular object (i.e., weight), (b) on what basis the value has been accorded (i.e., scale), and (c) the individual who is measuring the value (i.e., weigher). Knowledge of these three features might well be conditions for providing an initial characterization, although we might think that we would need to have examples of (a) and (b) on a number of occasions, (cf. BGE 263) to come to a more comprehensive understanding on the taste in question as an affective-evaluative sensibility or ‘*pattern* of conscious affectivity’ (Poellner 2007: 231, my emphasis). Naturally moving from an *understanding* of the taste in question to an *evaluation* of that taste (i.e., whether it is a ‘good’ or ‘bad’, ‘higher’ or ‘lower’, ‘base’ or ‘noble’), and what criteria *that* evaluation is being made on is a separate question, one which will attempt to deal with in Chapter 3 Section 4.

⁶³ Leiter thinks Nietzsche has first-order evaluative preferences, his ‘taste’, but that he seeks to provide no justification for them or impel anyone who does not already share them to adopt them (see Leiter 2000: 277-297). I will discuss this anti-realist reading of Nietzsche’s meta-ethics in Chapter 3.

⁶⁴ Harcourt 2011: 271.

3. Normativity and Exemplars

3.1 Preliminary Remarks

At the beginning of this Introduction I suggested that examining Nietzsche exemplars, through which I suggest he provides his evaluative ideals, is one way of understanding his philosophy of value. In Section 1 I considered the problem of prescriptivism vs. particularism in relation to Nietzsche and then in Section 2 considered some of the responses to this dichotomy, seeing how the two standard lines of interpretation, self-eliminativism and formalism, are unsatisfactory. I am now in a position to provide some more detail on the notion of exemplarity and explain how it might be able to overcome some of the difficulties that have been encountered so far.

At the end of criticizing the formal approaches I noted the importance of the qualitative dimension to Nietzsche's evaluative practices that both gives his philosophy of value substance and makes prescriptivism, and relatedly the question of the grounds he could provide for his evaluative preferences, a problem that needs to be addressed. The idea of exemplars representing Nietzsche's evaluative ideals brings this problem back into view, since we might think that exemplars are, in at least some sense, there to be followed. In this way, the appeal to exemplars might seem merely like a more idiosyncratic way of providing an ethics which claims to know what Zarathustra claimed not to know; 'the way' (Z, 'On the Spirit of Gravity', 2). To understand how in suggesting we follow exemplars Nietzsche does not betray his opposition to prescriptivist ethics we need to consider in more detail the kind of 'following' he is opposed to and his justification for this. Doing so will then enable me to specify the space his exemplars can occupy. To do so I will turn to the idea of the death of God and its relation to moral authority.

3.2 The death of God and moral authority

Nietzsche introduces the idea of the death of God at various points (GS 108, 125, 343, 357), and often alongside a speculative historical explanation as to how the pursuit of a certain kind of truthfulness, putatively associated with the Christian God and Christian Morality, came to undermine that faith (GS 357). Here I want to focus on the significance of the idea in relation to moral authority. On a basic reading, the death of God represents the collapse of a specific kind of divinely authorized, unconditional moral authority, that was central to religious-moral prescriptions. These prescriptions simply lose their normative authority insofar as the deity from whom they issued is no longer believed in. This captures part of the significance of this event, however it seems Nietzsche also something else in mind. He says:

Even less may one suppose many to know at all what this event really means - and, now that this faith has been undermined, how much must collapse because it was built on this faith, leaned on it, had grown into it -for example, our entire European morality. (GS 343 cf. GS 125)

So, aside from spelling the end of religious metaphysics, and therefore the foundation for religious moral authority, Nietzsche seems to think this event has a broader significance. As Aaron Ridley notes, ‘Nietzsche wants to persuade his audience that God’s death changes everything: that we are no longer entitled to our accustomed frame of reference and points of orientation’.⁶⁵ So we could say that the death of God does not just undermine religious moral authority, but might also mark a shift in the way in

⁶⁵ Ridley 2007: 92.

which normative claims have authority over us. If this is right then recognizing the significance of the death of God would involve challenging, or perhaps more strongly changing, our ‘frame of reference and points of orientation’ in ethics.⁶⁶ Nietzsche certainly seems to imply something like this with his claim that it will have a destabilizing effect on our ‘entire European morality’ (GS 343). However, we need to spell out in more detail what exactly this traditional frame of reference amounts to with regard to moral authority, and why the death of God should be taken to undermine moral authority *tout court* rather than just religious moral authority.⁶⁷

Presumably before mono-theistic faith became questionable morality had, at least predominantly, taken the form of externally imposed laws which, as was noted in my definition of prescriptivist ethics (Section 1), were taken apply categorically to all persons regardless of whether or not they were aiming to be moral. This provided humanity with what Nietzsche calls a ‘universally recognized goal’ (D 108), and it is the apparent impossibility of sustaining this kind of normative aim that the death of God signifies. As Ridley claims:

Before that event, mankind did possess ‘a universally recognized *goal*’, namely, salvation, in pursuit of which it was no at all ‘irrelevant and trivial’ to follow and act upon (divinely sanctioned) prescriptions and commands, that is, an externally imposed ‘moral law’. But prescriptions and commands require a prescriber or commander; and, with God’s death, that is precisely what goes missing. (Ridley 2007: 95)

⁶⁶ Ibid.

⁶⁷ A limited way of reading the death of God would be as only undermining the specific teachings of Christian Morality, and this is an obvious consequence of such disbelief as Nietzsche notes at TI IX 5. However, as will become clear in the discussion I think the import is more complex.

If this is right then Nietzsche thinks that the death of God undermines the traditional model of moral authority insofar as it brings to an end the normative warrant for ‘universally recognized goals’, and so also the project of prescriptivist ethics which, as we noted in Section 1, takes the form of providing categorical prescriptions to all persons. In this sense we might say that with the traditional source of moral authority ‘gone missing’ we do not merely lose one particular, mono-theistic, external commander or prescriber, but more broadly the warrant for following any externally imposed commands and prescriptions.

However, this reasoning depends on a concealed premise. For the conclusion to follow we would also have to accept that any justification for a ‘universally recognized goal’ (D 108), and so a justified source of moral authority of this kind (as giving rise to externally imposed, commands and prescriptions), can *only* be based on the existence of a theistic, or suitably transcendent, commander or prescriber. Note that it will not suffice to show that whatever exactly ‘European Morality’ is, it grew out of theistic presuppositions (GS 343). Since even if such a claim could be substantiated,⁶⁸ it may well be the case that in the course of its history, a traditional prescriptivist ethics has gained a rational warrant on separate, non-theistic, grounds, and that it therefore possesses an independent, justified source of normative authority for its externally imposed prescriptions.

What Nietzsche would therefore need to justify the conclusion that the death of God undermines all prescriptivist ethics, and their claims to moral authority, is something like the following premise:

- (1) All prescriptivist ethics are based, at the fundamental level of normative authority, on a theistic presupposition, even if concealed.

⁶⁸ Nietzsche seems to believe it can be, with the first essay on the *Genealogy* going to some way to providing an argument for this point.

However, the prospects for showing this premise to be even *prima facie* plausible seem poor. Since whilst Nietzsche claims to have grounds for something like this claim with regard to Kantian ethics, in terms of what he takes to be its ‘theistic’ appeal to noumenal realms and purely rational subjectivity (GM III 12, A11), in the case of, for example, Utilitarianism, or the political ideologies Nietzsche associates with ‘European Morality’ (such as socialism and egalitarianism), it is more difficult to see the concealed theistic presupposition. If, as Nietzsche sometimes suggests, these forms of prescriptive ethics (if that is indeed what they are) are dependent on a premise, say belief in equal rights, which putatively has its origins in some theistic idea, say equality before God (A 62, WTP 765), then the possibility of an independent rational warrant for such a belief can be put forward.

Given the difficulty of proving grounds for (1) I think the best way to interpret the impact that the death of God has on questions of moral authority is (a) as undermining one central source of traditional moral authority, namely theistic moral authority, and (b) having the additional effect of placing the burden of rational justification on those who wish to sustain the traditional model of moral authority. In this sense whilst Nietzsche would need a stronger argument to show that this kind of moral authority is impossible *per se* in virtue of the death of God, this event does have more significance than merely undermining the religious basis for morality.⁶⁹ It effects a more sceptical attitude towards the project of prescriptivist ethics, focusing attention on the fundamental question that such ethics will need to answer, namely what justification is being given for their categorical claims. So, whilst Nietzsche views the project of trying to ground traditional moral authority as an ‘erudite form of good *faith* in the dominant

⁶⁹ Simon Robertson makes the same point; ‘while Nietzsche is sceptical that *any* attempt (secular or otherwise) to ‘furnish the *rational ground* of morality (BGE 186) will succeed, he should not (and does not always) deny the *intelligibility* of such attempts’ (Robertson 2012: 95).

morality' and as such 'already situated with the terms of a certain morality' (BGE 186), we would nevertheless have to judge each attempt on a case-by-case basis.

3.3 Exemplary Ethics: prescribing and recommending

I noted above that Nietzsche is plausibly read as thinking that there has been a shift in the way in which normative claims have authority over us, such that the burden is now on those who think that traditional moral authority can be rationally justified to provide the argumentation. So, if we grant Nietzsche's scepticism about the possibility of a rational warrant for a prescriptivist ethics, alongside the fact that he does not want to provide *this* kind of ethics, then we should ask what alternative position his own evaluative ideals occupy.

If we look to Christine Korsgaard she sets out the normative landscape as follows:

Ethical standards are *normative*. They do not merely *describe* a way in which we in fact regulate our conduct. They make *claims* on us; they command, oblige, recommend, or guide. Or at least, when we invoke them, we make claims on one another. When I say that an action is right I am saying that you ought to *do* it; when I say that something is good I am recommending it as worthy of your choice. (Korsgaard 1996: 8-9)

However, for Nietzsche at least, this picture is ambiguous and does not make enough of a distinction between the traditional model of moral authority we considered in Section 1 (and the previous section), i.e. prescriptivist ethics, and a potentially revised normative framework within which ethics only recommends or guides. And it seems like it is this latter normative space that Nietzsche thinks ethics is still entitled to occupy:

To recommend a goal to mankind is something quite different: the goal is then thought of as something which lies in our own discretion; supposing the recommendation appealed to mankind, it could in pursuit of it also impose upon itself a moral law, likewise at its own discretion. But up to now the moral law has been supposed to stand above our own likes and dislikes: one did not want actually to impose this law upon oneself, one wanted to take it from somewhere or discover it somewhere or have it commanded to one from somewhere. (D 108)

Some of the ideas in this passage are complex enough that I shall hold off on discussion of them until the relevant sections of the main chapters of the thesis, specifically the way in which a ‘moral law’ might be in tune with one’s motivational repertoire in way that prescriptivist ethics did not assign importance to.⁷⁰ However, with regard to the broader normative picture, the meaning of the passage is clear enough. It suggests that at the level of methodology Nietzsche would have no general objection to an ethics that merely recommended certain ideals. And the idea of exemplars, through which Nietzsche could recommend his own evaluative ideals, seems tailor made to fit this space. As Ridley remarks, Nietzsche’s exemplars would, ‘provide an *occasion* for an audience to reflect upon itself (by comparing itself to the exemplary figure with which it has been presented) and then, at least, potentially, to live differently, or to think about itself differently, as a result’.⁷¹ In this sense an ethics based around evaluative ideals as presented through exemplars would not be prescriptive.

⁷⁰ It is a further complex question whether or not Nietzsche retains certain deontic notions, such as ‘ought’ or ‘obligation’, and if he does what import they might have. I will deal with meta-ethical issues in Chapter 3, although I am inclined to agree with Robertson who claims that rather than jettisoning deontic concepts altogether Nietzsche’s aim is ‘rather to question whether, absent an external law-giver, normative claims deploying those concepts have the normative authority traditionally supposed. Indeed, Nietzsche continues to frame his own positive claims via deontic notions’ (Robertson 2012: fn. 17, 95-96).

⁷¹ Ridley 2007: 101.

Yet this does not mean that one cannot learn a great deal from these exemplars. It just means that Nietzsche's ethics, as so understood, is not a form of moral pedagogy directed categorically at all persons from some external position of moral authority.⁷²

As Harcourt notes:

There seems to be plenty of room to hold an ideal – a conception of how to live such that one can say what's good about it – without any implication that other people are required to live according to it. Indeed one might think this is what an ideal *is*: a conception of how to live well that goes beyond what is required of one. (Harcourt 2011: 272)

So, as D 108 highlights part of what 'lies in our own discretion' is the degree to which we learn from the kinds of exemplars, as models of evaluative practice, that Nietzsche presents us with. And, in line with Nietzsche's particularist commitments we can see that learning is different from imitating, critical engagement is different from discipleship, and recommending or guiding is different from providing externally prescribed commands. Ridley therefore is right to conclude that:

The offering of exemplars...lines up firmly on the recommending side of the distinction between recommending and commanding – and so, in consequence falls naturally into just the sort of territory that Nietzsche needs to occupy if he is to pursue his project within the constraints that concern him. (Ridley: 2007: 101)

So, it is reasonable to think that Nietzsche's use of exemplars can provide a potential resolution to the dichotomy between prescriptivism on the one hand, and the problems we saw emerge with formal approaches to his ideal on the other. Since, through his exemplars Nietzsche can and does recommend contentful ideals, as models of

⁷² See D 194.

evaluative practice, that are concerned with what we value and how we value. Yet, this is a legitimate practice for him insofar as (a) these recommended models are not categorically prescribed and (b) if reflected upon, should point back to something internal to the individual who seeks to learn from them, and who by applying them in his own case would stand to achieve something distinctly for himself in the way Nietzsche has Zarathustra suggest (Z, ‘On the Spirit of Gravity’, 2).

Concluding Remarks

Hopefully this final section has provided at least a *prima facie* rationale for the methodology that will be pursued in the main chapters of my thesis. However, to be clear about what has been suggested here I want to reflect on a point about my concern with Nietzsche’s first-order evaluative practices.

All this final section has shown is that Nietzsche can, without obvious inconsistency, provide a philosophy of value with a normative dimension, where the latter is based, at the level of methodology, on making recommendations rather than prescriptions. What it has *not* provided is either (1) the actual substance or content of Nietzsche’s ideal in terms of evaluative ideals or (2) any comprehensive account of Nietzsche’s grounds for those particular evaluative preferences as to be given in those evaluative ideals. It is the aim of the main chapters to provide a comprehensive picture of (1) through consideration of three central exemplars. However, in doing so, and with the aim of presenting a coherent picture of Nietzsche’s philosophy of value, (2) will also be given due consideration. So the hope is that by looking at Nietzsche’s exemplars as representing his evaluative ideals I will be able to provide some key insights on the central issues and problems of his philosophy of value.

I

The Critical Ideal: *The Free Spirit as Exemplar*

The aim of this chapter is to provide an account of the exemplar of the *free spirit* as found primarily in *Human, All Too Human* (hereafter HA). The critical ideal that this exemplar represents will be defined through liberation from the ‘metaphysical need’, where this represents a ‘corrupted’ affective-evaluative sensibility Nietzsche identifies as present in metaphysical philosophy and certain ‘moral sensations’. Understanding the *free spirit* in terms of this critical ideal is therefore the first step towards the more positive philosophy of value that will be developed in the subsequent chapters of my thesis. Section 1 will provide an overview of Nietzsche’s methodology in HA. Section 2 will consider his critique of metaphysics and the evaluative attitude of metaphysical indifference.¹ Section 3 will examine Nietzsche’s rejection of certain notions of free will and explain how this relates to questions of moral responsibility. Finally section 4 will consider Nietzsche’s account of guilt.

1. Nietzsche’s Methodology in HA

In its original publication HA opened with a passage taken from Descartes ‘*Discourse on Method*’,² yet Nietzsche never held specific Cartesian views.³ Rather what he takes from Descartes is a sceptical attitude which involves questioning firmly held, yet

¹ This term is taken from Poellner 2009: 170-177, also Poellner 2001: 111-117.

² Descartes 1998.

³ As we shall see Nietzsche’s criticisms of metaphysical philosophy would apply to Cartesian immaterial substances and the kind of *a priori* reasoning Descartes uses to establish his metaphysical claims.

potentially rationally suspect, beliefs. As Descartes puts it, he ‘learned not to believe anything too firmly of which I had been persuaded only by example and custom’ and freed himself from ‘many errors that can darken our natural light and render us less able to listen to reason’.⁴ However, if this were all Nietzsche’s methodology in HA amounted to then we might struggle to see its distinctiveness from its predecessors (for example in Hume or Kant). Perhaps though Nietzsche could argue that this sceptical attitude was, in certain instances, not carried far enough,⁵ or was suspended in order to sustain ‘motivated’ beliefs. Such beliefs are not held, at least not primarily, due to their veridicality or explanatory power, i.e., for what we might think of as good reasons, but due to their importance in supporting certain cultural and evaluative practices, i.e., for motivated reasons. For example, Nietzsche often suggests that belief in contra-causal free will, that is, the idea that human beings have the capacity to step outside the causal order and make choices which originate exclusively from their own will, is motivated, in part, by the necessary role that such a conception plays in supporting the practice of retributive justice (GM I 13, cf. HA 105, and Section 3 of this chapter). Uncovering such motivated beliefs, as we will see throughout this chapter, is a key part of the critical ideal that the *free spirit* exemplifies.

However, Nietzsche also seems committed to a more specific methodology which he calls ‘historical philosophy’ (HA 1-3). Whatever this historical philosophy involves Nietzsche thinks it is different from what he labels metaphysical philosophy, which he accuses of having avoided ‘questions concerning origins and beginnings’ (HA 1), and lacking ‘historical sensibility’ (HA 2).⁶ These ‘failings’ are of no small measure insofar

⁴ Descartes 1998: 6. Matthew Meyer makes the implausible claim that in HA Nietzsche in fact adopts ‘the rationalism of Descartes’ (see Meyer 2004: 7).

⁵ In the case of Hume this claim would be difficult to establish, since alongside Nietzsche, Hume represents as thoroughgoing a sceptical attitude as can be found in modern philosophy.

⁶ Presumably the main targets Nietzsche has in mind here are Schopenhauer and Kant. See

as Nietzsche goes on to claim that they lead to a mistaken conception of the nature of human beings (I will return to this point in more detail below). Sections 2, 3 and 4 of this chapter will deal with Nietzsche's critique of metaphysical philosophy and its moral-religious counterparts. What I want to focus on first are the presuppositions of Nietzsche's own historical philosophy.

Nietzsche describes his historical philosophy as a 'philosophical science' (HA 1) and tells us that we should not think of this 'youngest of all philosophical methods' as 'separate from natural science' (HA 1). However, the term Nietzsche uses for science, *Wissenschaft*, does not just denote the physical sciences (i.e., physics, chemistry) but implies a wider reference to any rigorous discipline which takes an observational and broadly empirical approach to studying its object(s). In this sense we might say any 'science' typically employs an *experiential* method, that is, testing and revising its claims against observed phenomena, leading it to propose more accurate and finely-grained hypotheses, such that the relevant knowledge might be said reveal itself 'only to the keenest observation' (HA 1) through the application of 'rigorous methods' (HA 3, cf. HA 633, A 59).⁷ This might seem a relatively non-specific method for philosophy to pursue, yet, it rules out the 'first philosophy' of Descartes and the Rationalists (e.g., Spinoza, Leibniz, and also earlier religious-metaphysical thinkers such as Aquinas, Anselm), who attempt to establish at least some of their philosophical claims through *a priori* reasoning.

Moreover, the connection with 'natural' science seems to suggest that Nietzsche's philosophy in HA is in some sense naturalistic. To unpack what this might amount to Nietzsche's claim to be providing a 'chemistry of concepts and sensations' (HA 1) is

Schopenhauer 1969a: 273-4, and Kant 1991: 58.

⁷ As Leiter puts it, *Wissenschaft* for Nietzsche means 'any rigorous methods of inquiry, whose methods secure the epistemic reliability of their results' (Leiter 2015 fn.11: 6).

partly revealing, insofar he claims his philosophy is searching for what he calls the ‘fundamental element’ (HA 1), although I shall argue below that the allusion to chemistry should not to be taken too literally. In keeping with the idea of a philosophical *Wissenschaft*, the search for a fundamental element might involve the positing of some causal explanation for a pattern of observed phenomena, like the way a chemist might posit a heat source as an explanation for observed steam. So, when Nietzsche praises scientific methods and calls for a ‘chemistry of moral, religious, aesthetic representations and sensations’ (HA 1) it seems reasonable to assume that he is at least committing to some kind of *methodological* naturalism, one which involves positing causal explanations. However, understanding what, other than applying an experiential method, such a commitment to naturalism entails is more difficult.

Leiter claims that Nietzsche is a methodological naturalist in the sense that he thinks that ‘philosophical inquiry...should be continuous with empirical inquiry in the sciences’, although much depends on how this continuity is cashed out.⁸ Drawing on Nietzsche’s use of the term *Wissenschaft*, we should note that this kind of methodological naturalism might just as well imply continuity with the social sciences as with the physical sciences. This is an important qualification insofar as it allows Nietzsche to be a methodological naturalist, in the sense of applying some kind of experiential method in philosophy, without being committed to naturalistic reductivism, as found in contemporary physicalism, functionalism, and the German Materialists of Nietzsche’s own time.⁹ Naturalist philosophy of that stripe is typically committed to strict continuity with the physical sciences and so the additional claim that it is only the

⁸ Leiter 2015: 2. For an alternative of what Nietzsche’s naturalism might amount to see Janaway 2007: 34-53. For Leiter’s response see 2015: 244-252.

⁹ On Nietzsche’s connection to the German Materialists see Leiter 2015: 50-56.

properties picked out by those sciences, on some extension of them, that are in fact real, and therefore the only legitimate materials of causal explanations.¹⁰

Nietzsche's request for a 'chemistry of moral, religious, aesthetic representations and sensations' (HA 1) could be misleading with regard to the above, since if read literally it might seem committed to precisely this kind of reductivism, such that a moral sensation, for example, is to be reduced to brain chemistry. Yet, careful attention Nietzsche's philosophy, both in HA and elsewhere, undermines the idea that he holds this position. His persistent appeal to a whole range of psychological causes in explanations for human phenomena, rather than physical states or chemicals in the brain, and his claim that 'psychology again be recognized as queen of the sciences, and that the rest of the sciences exist to serve and prepare for it' (BGE 23), rules out such a reading.¹¹

In this sense the 'fundamental element' that 'reveals its presence only to the keenest observation' (HA 1), that is, the causal elements figuring in the kinds of explanations Nietzsche offers for the kinds of complex, often evaluative, phenomena he is interested in (such as metaphysics and morality), would, by his own lights, not be satisfactorily accounted for in terms of those elements picked out by the physical sciences.¹² Rather,

¹⁰ Physicalism about the mind usually comes with the claim that conscious-states are epiphenomenal and causally inert. Although there is a difference between 'reductivism' in the sense of holding that conscious-state *a* is *reducible*, and so to be causally explained in terms of, say, brain-state *b*, and *eliminativism* which would hold that since conscious states cannot be satisfactorily explained in terms of a certain naturalist framework then they should be eliminated from our philosophical understanding.

¹¹ Leiter makes the same point; 'there is no evidence that Nietzsche is at all sympathetic to this latter kind of S-naturalism', i.e., to a kind of *reductive* naturalism (Leiter 2015: 4-5). For other passages which show Nietzsche's hostility to *reductive* naturalism, usually framed in terms of his opposition to 'materialism' and 'mechanism' see GS 373, GM III 16, and BGE 12.

¹² There is more complex question concerning the degree to which psychological states supervene on brain states or brain chemistry. I do not have the space to consider such questions

the explanatory priority and indispensability of psychological concepts and explanations, which often appeal to the role that *qualitative* (or phenomenological) features play in identifying and distinguishing certain kinds of conscious states, seems to be one of the most recurrent features of Nietzsche's philosophy in HA and elsewhere.¹³ Such considerations should count against too literal a reading of Nietzsche allusion to chemistry in HA 1, and the reference should point back to his commitment to a kind of *methodological* naturalism, specifically the application of an experiential method whereby this implies the postulation of, and search for, observable and verifiable causes. As Nietzsche remarks later in HA, 'on the whole, the *methods* of science are at least as important an outcome of inquiry as any other result' (HA 635, emphasis mine cf. HA 256).

However, there does seem to be one substantively naturalistic claim that Nietzsche is committed to, which we might identify as a kind of ontological-naturalism, or as Leiter puts it, the 'view that the only things that exist are *natural* things'.¹⁴ One way of understanding this aspect of Nietzsche's method might be as a thoroughgoing anti-super-naturalism, such that he allows no role for super-natural entities (e.g., God, 'things-in-themselves') in explanations of the phenomena he is interested in. This is part of the story, although we can say more. Insofar as ontological naturalism is an underlying assumption of a philosophy, we might think that the explanations that it offers should not posit causes or capacities which are either in principle undiscoverable

here, and they are in any case tangential to Nietzsche's concerns.

¹³ Paradigmatic cases are his description of Nietzsche's description of *ressentiment* in GM I, and nihilism in WTP 1-25. However, there is legitimate question to be asked about the role of Nietzsche's appeal to physiological causes to explain certain evaluative phenomena, and whether this might imply a kind of biological reductivism. For discussion of this feature and its relation to Nietzsche's method see Poellner 2006: 297-313.

¹⁴ Leiter 2015: 4. Leiter links this commitment back to Nietzsche's reading of the German Materialists.

by any rigorous (empirical) science or are of a fundamentally non-natural character (i.e., contra-causal free will, God, *Will* à la Schopenhauer). Rather, such a philosophy should provide explanations and descriptions of its objects of inquiry in terms that are broadly of a piece, and certainly not obviously opposed to, or radically different from, the kinds of explanations, properties, and causes picked out by rigorous science.

This certainly seems to be guiding heuristic that Bernard Williams draws from the critical dimension to Nietzsche's project, claiming with regard to moral psychological explanations, but reasonably generalizable to all philosophical explanations, that they should add as little as possible of what we might call 'domain specific' capacities, and that when they do 'we have to ask whether there may not be a more illuminating account that rests on conceptions that we use anyway elsewhere'.¹⁵ It seems like this description of Nietzsche's method would make most sense of his suggestion that philosophy should not be '*separate* from natural science' (HA 1, my emphasis), and certainly fits the way he accounts for his philosophical task later, as translating 'man back into nature' (BGE 230), whilst maintaining, as I think Williams correctly identifies, that 'what is at issue is not the application of an already defined scientific programme, but rather an informed interpretation of some human experiences and activities in relation to others'.¹⁶

So, Nietzsche's methodology in HA has two aspects so far; (1) exposing motivated beliefs, and (2) a kind of naturalism embracing the experiential method and ontological naturalism. However, neither of these features particularly explains him calling his philosophy 'historical'. However, there is a more general import to Nietzsche's use of this term, and one which plays into (1) and (2).

¹⁵ Williams 1993: 6.

¹⁶ Williams 1993: 7. See Leiter 2015: 3-6 for a more detailed account of the kind of naturalist Nietzsche might be.

Nietzsche claims that a central task of his historical philosophy involves restating questions of origins with what he calls ‘historical sense’, claiming that ‘lack of historical sense’ with regard to origins has led to a misunderstanding of what human beings, as of a piece with the natural world, fundamentally are (and their capabilities):

All philosophers have the common failing that they start with present-day human beings and suppose that they will reach their goal by analyzing them. Involuntarily, they allow “man” to hover before their eyes as an *aeterna veritas*, something that remains the same through all turmoil, a secure measure for things. (HA 2)

Historical philosophy is distinguished from at least some previous philosophy then in terms of its conception of the human being as something that is not a fixed ‘eternal truth’ but rather has ‘come to be’ and as such, is capable of being otherwise than it is at present. However, we should be careful in understanding what this ‘being otherwise’ amounts to, since Nietzsche should not be read as committed to a conception of human beings, often misassociated with certain brands of existentialism, as entirely protean, such that their nature is a matter of essence-determining radical choice. Rather, human beings for Nietzsche do have *some* essential qualities, as he states ‘everything *essential* in human development occurred during primeval times’ (HA 2), it is just that what metaphysical philosophy has claimed to be essential is, according to him, ‘nothing more than testimony about the human being of a *very restricted* stretch of time’ (HA 2). However, for Nietzsche, the ‘lack of historical sense’ of metaphysical philosophy does not lie merely in its ignorance of specific historical narratives (e.g., evolutionary ones), but in its claims about what human beings essential qualities or ‘nature’ amounts to, i.e. a metaphysical nature with certain capacities which ‘originate’, in some sense, outside the natural world, having what Nietzsche calls a ‘miraculous origin, immediately out of

the core and essence of the “thing in itself” (HA 1). We will see more concrete examples of these kinds of claims in the later sections of this chapter.

These observations, whilst making more sense of the historical dimension to Nietzsche’s methodology in HA, might not seem to add that much, although with some further reflection we can see how they might. As will become clear throughout this chapter the problem for Nietzsche with metaphysical philosophy and its religious-moral counterparts is not merely certain descriptive mistakes he thinks they make concerning qualities or capacities human beings are claimed to possess, but also the way they have cultivated something of a ‘corrupted’ taste in terms of the ‘metaphysical need’ (remember that, as discussed in my Introduction, taste should be understood in terms of a certain affective-evaluative sensibility; a pattern of conscious affectivity in terms of a characteristic way of responding and evaluating).

With this in mind a ‘liberating philosophical science’ (HA 1) might develop a certain kind of historical understanding both as a means to understanding (a) how human beings came to interpret themselves as possessing a certain nature or certain qualities, and also (b) how they came to develop a certain affective-evaluative sensibility. And it would offer this understanding as a means to potentially overcoming not just erroneous beliefs, but also overcoming a ‘taste’ that remains psychologically embedded even after certain metaphysical or religious beliefs are no longer, at least not explicitly, held as true. It is with this aim in mind that Nietzsche claims that after we have attained the negative goal ‘that every positive metaphysics is an error’ then we ‘must grasp the historical justification as well as the psychological one for such conceptions’ (HA 20).

2. Nietzsche's Critique of Metaphysics

2.1 Preliminary Remarks

The aim of this section is to examine Nietzsche's critique of metaphysical philosophy, primarily with reference to passages from HA. I will argue that Nietzsche's critique consists of two stages, (1) an attempt to 'refute' the possibility of a certain kind of metaphysics (which will be labelled 'rich metaphysics' or RM), and (2) the suggestion that we should be indifferent to the only type of metaphysics ('thin metaphysics' or TM) that remains after 'rich metaphysics' is 'refuted'. Given this two-fold aim, the details of Nietzsche's critique of metaphysics are complex and will depend on clarifying the different positions he takes on the respective targets. I will begin with some remarks on HA 9. There Nietzsche says:

It is true, there could be a metaphysical world; the absolute possibility of it can hardly be contested. We see all things through the human head and cannot cut this head off; and yet the question remains as to what part of the world would still be there if one had in fact cut it off...what has up to now made metaphysical assumptions *valuable, terrible, pleasurable* for them, what engendered such assumptions, is passion, error, and self-deception; the worst of all methods of knowledge, not the best of all, have taught us to believe in them. When we have revealed these methods to be the foundation of all existing religions and metaphysics, we have refuted them. That other possibility remains, but we cannot even begin to do anything with it, much less to allow happiness, salvation, and life to hang from the spider threads of such a possibility. – For we could assert nothing at all about the metaphysical world except its otherness; it would be a thing with negative characteristics. – Even if the

existence of such a world were to be proven so well, any knowledge of it would certainly still be the most irrelevant of all knowledge. (HA 9)

This passage contains a significant amount of material and I will return to a number of its claims throughout this section.

To begin, I want to emphasize that Nietzsche neither denies what he calls the ‘absolute possibility’ of a metaphysical world, nor some kind of knowledge pertaining to it. That is to say he is not proposing the very conception of a metaphysical world, and knowledge of it, to be, at least in theory, nonsensical or meaningless.¹⁷ Yet at the same time he tells us, presumably after his critique of metaphysics, that we will come to the position of seeing ‘all existing religions and metaphysics’ as having been ‘refuted’ (HA 9). These claims might seem *prima facie* incompatible, after all if metaphysics has been refuted this seems to undermine any sense to talking about metaphysical knowledge. Also we might wonder why, if metaphysical philosophy has been refuted, we should countenance, even as just a theoretical possibility, the existence of a metaphysical world.

However, to overcome these initial worries we should make a distinction between two different types of metaphysics being targeted in this passage. On the one hand we have what Nietzsche refers to as ‘metaphysical assumptions’ or elsewhere as ‘positive’ metaphysics (HA 20; more on this in the following section). On the other hand we have that ‘other possibility’, which according to Nietzsche primarily concerns theoretical

¹⁷ Although as Poellner notes Nietzsche does, elsewhere ‘express doubts about the availability of such metaphysical knowledge’ (Poellner 2009: fn.38). See Poellner 1995: 57-78, 150-61, 173-86 for a discussion of this dimension to Nietzsche’s later thinking on metaphysical knowledge. Poellner’s reading is however different from Clark’s interpretation of Nietzsche’s mature thought that she claims ‘repudiates the thing-in-itself as a contradiction in terms’ (Clark 1900: 102).

speculation about the ‘absolute possibility of a metaphysical world’ (HA 9). With these two different conceptions in mind we can specify the different attitudes Nietzsche takes towards each respectively.

Put simply, it is the former, positive metaphysics, that Nietzsche wants to refute, and as the end of the passage suggests, it is the ‘negative’ metaphysics which he claims we should, in some sense, be indifferent towards, that is to say, knowledge claims pertaining to the existence of a ‘metaphysical world’ in this second sense would, for reasons we have yet to consider, be of no cognitive interest to beings like us, they would provide only ‘the most irrelevant of all knowledge’ (HA 9). So, we can see that Nietzsche does not think that metaphysical knowledge or metaphysical assumptions are *per se* irrelevant, or that we should be indifferent to metaphysics *tout court*. Rather, as the desire to provide a refutation shows, Nietzsche is in fact very much concerned with the critical aim of showing that ‘every positive metaphysics is an error’ (HA 20). As Poellner notes, ‘Nietzsche’s own attitude towards metaphysics is not simply globally indifferent or quietist’.¹⁸ However, to move from Nietzsche’s initial critique of metaphysics to his final position, that we should be indifferent to a certain kind of metaphysical knowledge, requires understanding that first aspect. In other words, the logic of Nietzsche’s position is such that a certain kind of metaphysical knowledge will only become irrelevant in way he envisages if he is successful in showing that ‘positive’ metaphysics is indefensible.¹⁹ Yet, before considering Nietzsche’s argument for the

¹⁸ Poellner 2009: 175. By contrast a good example of such a globally quietist approach to metaphysics might be found in Rorty 1989: 8.

¹⁹ Clark glosses over this point and fails to see that Nietzsche’s critique of metaphysics is a more complex project than can be captured by a description of HA’s epistemology as a kind of epistemic agnosticism (Clark 1990: 99-102). Although 8 years later she amended her position, ‘the major project of *Human, all too Human* is to induce scepticism concerning the metaphysical world by showing it to be cognitively superfluous’ (Clark 1998: 49-53), although

first dimension to his critique, it will be helpful to say more about these distinctive types of metaphysics.

2.2. *Rich-metaphysics and Thin-metaphysics*

In order to understand more clearly the different types of metaphysics in Nietzsche's critique we should make a distinction between what I will call rich metaphysics (hereafter RM) and thin metaphysics (hereafter TM). As a first attempt at defining RM we might say something like the following:

(1) *The content claim*: RM posits a metaphysical world about which it claims to have substantial knowledge in terms of positive characteristics.

However, this is a relatively sparse characterization, and in order to be appropriate to Nietzsche's discussion of metaphysical philosophy in HA 9, and elsewhere, we need to build in a number of caveats. So we might make the following amended claim:

(1*) *The amended content claim*: RM posits a 'real' metaphysical world about which it claims to have substantial knowledge in terms of positive characteristics, characteristics that are *not* those of the 'apparent' world we experience.

Something like (1*) captures several of the actual metaphysical theories and philosophies that are likely the target of this first dimension of Nietzsche's critique, e.g., Schopenhauer's metaphysics of *Will* and *Representation*, Kant's (on one reading at least) distinction between noumenal and phenomenal reality, the ideal forms of Plato,

again it seems the relevant distinctions between *types* of metaphysics are not made explicit.

and the transcendent realm of Christian metaphysics. What all these have in common is the positing, in some form, of a ‘second’ or ‘real’ metaphysical world, about which positive claims are proffered, and which is asserted to be different in its fundamental characteristics from the ‘apparent’ world of ordinary experience. In a later work Nietzsche talks in these terms of those, specifically Kant and Christianity, who ‘divide the world into a ‘true’ half and an ‘illusory’ one’ (TI III 6).²⁰

However, even this formulation (1*) is potentially ambiguous since it could be thought to involve a disconnection between the ‘real’ metaphysical world and the world of ordinary experience. Since, if the ‘apparent’ world we experience, almost by definition, exhausts our actual and possible experience, then we might wonder by what means RM could establish their ‘positive’ claims, other than by *a priori* reasoning, and why any such claims, divorced from experience, should have any cognitive interest for us. In this sense if (1*) is all RM amounts to then Nietzsche’s suggestion in HA 9 that we should be indifferent to a certain kind of negative metaphysics seems *prima facie* just as applicable to positive ‘metaphysical assumptions’. However, not only would such a view undermine the distinctions Nietzsche is making in HA 9 between the different stances he thinks we should take on different types of metaphysics, but it would also involve a mischaracterization of at least some of those RM which are plausibly Nietzsche’s targets.

With regard to the latter point Schopenhauer’s characterization of the remit of metaphysics is revealing. He tells us that:

²⁰ This distinction is present in a large number of passages in which Nietzsche discusses metaphysical philosophy. See BGE 10, A 10, EH P 2, TI III 2, TI ‘How the “True World” Became a Fable’, EH IV 8, WTP 461, 507, 566, 567, 574, 579, 583, 586.

It [metaphysics] never really goes beyond experience, but only discloses the true understanding of the world which lies before it in experience. It is neither, according to the definition of metaphysics which even Kant repeats, a science of mere conceptions, nor is it a system of deductions from *a priori* principles, the uselessness of which for the *end* of metaphysics has been shown by Kant. But it is the rational knowledge, drawn from the perception of the external actual world and the information which the most intimate fact of self-consciousness affords us concerning it, deposited in distinct conceptions. It is accordingly the science of experience. (Schopenhauer 1969b: 183)

What this passage makes clear is that far from there being a disconnection between the ‘true understanding of the world’, i.e. the ‘real’ metaphysical world of RM, and our experience of the ‘apparent world’ (or as Schopenhauer puts it, ‘the perception of the external actual world’), metaphysics, at least of this kind, seeks to establish a relation of some sort between the two. In this sense, as Schopenhauer makes clear, metaphysics proper, at least for him, does not speculate on the basis of *a priori* reasoning about a ‘real metaphysical world’, which divorced from actual experience might then all too easily be dismissed as irrelevant. Rather, in Schopenhauer’s case at least, the arguments for metaphysical claims are based on a certain interpretation of experience. For example, his own metaphysics is offered as a kind of inference to best explanation on the basis of certain, as he puts it, ‘immediately’ experienced facts of ‘self-consciousness’, those being experiences of a kind of first-personal bodily agency *as Will* which is then extended to, and claimed to be the ground of, all material reality.²¹

So, given these considerations, we might supplement (1*) with an additional claim:

²¹ Schopenhauer 1969a: 105.

(2) *The substantive relation claim*: RM present arguments for accepting certain metaphysical claims about the ‘real world’ with recourse to features identified by an interpretation of the ‘phenomenal world’ we do experience.

However, again matters might be more complex than (2) makes out. Whilst it certainly captures Schopenhauer’s metaphysical project it seems less applicable to the variety of ways in which positive metaphysical claims have been argued for in the history of philosophy, and as such threatens to rule out too many ‘metaphysical assumptions’ that are presumably also the target of Nietzsche’s critique. As Schopenhauer himself notes, his definition is distinct from Kant’s so-called ‘science of mere conceptions’.²² In order to capture a broader range of metaphysical views we need to introduce a caveat to (2) which will give us something approaching, along with (1*), a sufficient characterization of RM.

In order to see the amendment that needs to be made we can turn to the example of Christian metaphysics, since this will reveal the sense in which (2) is part of the story, but not the whole picture. In line with (2) Christian metaphysics undoubtedly argues for its own interpretation of our everyday experience. Traditionally certain forms of divine intervention are put forward for explaining a range of phenomena, arguing for what are sometimes called the observed effects of God in the world through things like miracles. In this sense (2) above certainly fits an aspect of Christian metaphysics. Yet, as found in religious philosophers such as Anselm and Aquinas, actual Christian metaphysicians also attempt to provide *a priori* arguments for theistic doctrines, such as the Ontological argument or Cosmological argument.²³ These arguments are not based on an

²² Schopenhauer 1969b: 183.

²³ See Aquinas 1993: 251-270, for the Cosmological argument and Anselm 1998: 82-105 for the Ontological argument. For a detailed critique of these arguments see Mackie 1982: 41-55

interpretation of experience (or at least not obviously), but nonetheless if true are relevant to our practical engagement with the world. For example, if it turned out that the Ontological argument was valid, and that, for rational subjects at least, God's existence was incontestable, Christian Morality's moral prescriptions would take on a more serious character. In this sense a substantive relation between the world we experience and the 'real' metaphysical world need not be based on Schopenhauerean-style arguments drawn from aspects of experience. Rather, such a relation might also emerge through the way in which the truth of certain metaphysical claims, regardless of how they are established or argued for, affects our attitude to the lives we live in the world we do experience.

So, given these reflections we might make a disjunctive amendment to (2):

(2*) *The amended substantive relation claim:* RM's *either* present arguments for accepting certain metaphysical claims about the 'real world' with recourse to features identified by an interpretation of the 'phenomenal world' we do experience *or* argue that the truth-status of certain metaphysical claims has an incontestable practical relevance for our 'phenomenal' world.

In this sense we might say that RM will always seek to establish some kind of substantive relation between the world we experience, and the 'true' metaphysical world (or 'true' metaphysical claims), although the nature of this relation need not be the same for different kinds of RM. Nevertheless, in both the ways suggested in (2*), we cannot avoid the fact that the claims being made by RM matter to us, they are not, or at least not obviously, 'irrelevant' (HA 9).

Pausing for a moment to reflect on the above amendment we can see that (2*) allows for a broader range of metaphysical philosophy to be the target of Nietzsche's critique of RM, since RM can still involve a theoretical or *a priori* dimension, in which, as Kant puts it, 'abstraction is made from all experience',²⁴ and in certain cases quite deliberately (e.g., Descartes). And this method can be used to argue for certain positive claims about a 'true' metaphysical world as different from the world of appearance (1*). This kind of method is found in the Rationalist philosophers (e.g., Descartes, Leibniz and Spinoza). However what (2*) makes clear is that for a metaphysics to count as instance of RM, and so be a target for the first stage of Nietzsche's critique, it must supplement its theoretical claims with either an implicit or explicit claim to a substantial relation with the world we experience. The most obvious example, as we saw above, are the practical consequences that are taken to follow from the rational proofs for the existence of God. Taken together then (1*) and (2*) are jointly sufficient to characterize a RM insofar as no metaphysics could exemplify these conditions and fail to be a target for the first stage of Nietzsche's critique.

However, before moving on to Nietzsche's so-called 'refutation' of RM something should be said about that 'other possibility', what I called thin metaphysics (TM). In the same way as I, in part, defined RM in terms of its variety of ways of establishing a substantive relation to world we experience, we might define TM by the absence of such a relation. That is to say a TM does not involve claim (2*); it does not base its claims upon empirical or phenomenal experience and does not seek to supplement its metaphysics with any substantive claims about the practical relevance of its claims. However, this should not be taken to mean that TM are entirely devoid of content, since, if there was no content at all to TM then an attitude of indifference would not represent anything particularly distinctive, there would quite literally be nothing to care

²⁴ Kant 1998a: A591/B619.

about. Poellner gives a good summary of examples of TM in the history of philosophy that fits this characterization:

While certain metaphysical theories thus have an evident and incontestable practical dimension, there are others that do not. Among these purely theoretical claims are many that historical and contemporary metaphysics have been, and continue to be, preoccupied by. They include the following familiar staples; that the real spatial world consists of absolute, non-perspectival objects (Descartes, Locke); that there are no non-perspectival spatial objects (Berkeley, Kant, sometimes Nietzsche); that the world does not include phenomenal consciousness among its ultimate realities (physicalism); that the real world consist only of consciousness and its objects (idealism). (Poellner 2009: 176).

I want to argue that it is this kind of purely theoretical metaphysics that Nietzsche wants, in the end, to suggest we should be indifferent towards, although as we shall see when we consider his metaphysical indifference in more detail the precise nature of this position is ambiguous. For now we can hold off on this detail and consider the first stage of Nietzsche's critique of metaphysics. We have a good sense of his target in terms of our characterization RM and his stated goal, namely to refute RM and show that 'every positive metaphysics is an error', i.e., that every RM is an error (HA 20). As we shall see Nietzsche's criticisms are mostly based on undermining (2*), that is RM claim to a substantive relation between its metaphysics and the world we experience.

2.3 Nietzsche's critique of RM

I argued above that RM, as the target of the first stage of Nietzsche's critique, represents a complex family of theories that are by no means uniform in the way they seek to

establish a relation between their metaphysical claims and experience. A critique of RM is therefore likely to be difficult, and Nietzsche takes on a considerable argumentative burden in claiming to be able to refute RM and show that ‘every positive metaphysics is an error’ (HA 20). Before considering how he attempts to do this, we should note a method of critique that would not be sufficient for Nietzsche’s critical purposes as he presents them here, that of *a priori* argumentation.

We saw in our introductory remarks on the experiential method that Nietzsche’s ‘historical philosophy’ generally opposes such ‘first philosophy’. Moreover, as was noted above, purely *a priori* or theoretical reasoning is also the purview of TM, and so we might think that if Nietzsche were to critique RM by such methods alone then he would stand to miss what makes the former distinct from the latter. For example, an exclusively *a priori* argument purporting to show that Schopenhauer’s metaphysics of the *Will* is logically fallacious or self-contradictory might well be successful, and this would surely count against this particular metaphysics.²⁵ Yet this method would not, at least not directly, undermine that substantive relation that RM seeks to establish (see 2* above), and given corrective dialectical moves on the part of its defender a particular RM might come away unscathed. In this sense we might think that Nietzsche’s critique of RM, given both its ambitions and the nature of its target, has to avail itself of tools other than, or as well as those of, exposing logical inconsistency through theoretical reasoning.

According to HA 9, what might bring about such a refutation of RM is a revealing of its methods as based in ‘passion, error, and self-deception’. Nietzsche tells us that ‘when we have revealed these methods to be the foundation of all existing religions and metaphysics, we have refuted them’ (HA 9). Such a claim seems problematic if by

²⁵ In fact Nietzsche provides us with something like this in his early-unpublished essay *On Schopenhauer*.

refutation Nietzsche means showing some claim to be *in fact* false, since the revealing of flawed methods is, strictly speaking, irrelevant with regard to the truth-status of the claim that is being made. The method for acquiring a belief and the truth-status of a belief do not stand in this kind of entailment relation. For example, guessing answers to multiple-choice questions is no doubt a flawed methodology if my aim to get the questions right, but this fact alone does not preclude the possibility that in certain instances I nonetheless might get the answers right.

However, a more charitable reading of Nietzsche's claim to be 'refuting' RM can be offered by interpreting refutation more modestly as the engendering of suspicion. Nietzsche in fact tells us that 'if we *mistrust* metaphysics, this has by and large the same consequences as if it had been directly refuted and we were no longer *allowed* to believe in it' (HA 21, first emphasis mine). In this sense Nietzsche wants to argue that revealing a flawed methodology will have the effect of bringing about a feeling of mistrust towards RM. However, we might again accuse Nietzsche of exaggerating his point, and question whether in fact mistrust no longer allows us to believe in something. For example, a Christian who has a moment of doubt in his religious beliefs can later re-affirm his faith. Yet, surely Nietzsche's point is that (a) we cannot mistrust *p* and sincerely, or without question, believe *p* at the same time, and that (b) if mistrust at least brings to conscious awareness the possibility of not-*p* then affirming *p* will require some proof otherwise, some good reasons to reject not-*p*. In the case of RM, then, Nietzsche wants to show that a defective methodology, which he purports to show lies at the basis of its knowledge claims, should make us mistrustful to the extent that our belief in RM, absent good arguments to the contrary, should be suspended; we should no longer be 'allowed to believe in it' in this sense (HA 21). However, the specific way in which this revealing of methods works in Nietzsche's critique is complex, so I will attempt to illuminate his position with an example.

Say an individual tells me that he believes in ghosts, yet I come to find out from close friends of his that he recently suffered a bereavement, and that he has since spent much of his time reading super-natural fiction, and has taken up a new occupation supervising ghost walks. Moreover, I am told by my confidants that prior to the bereavement, and his new occupation, he never mentioned ghosts or read such literature, but now is explicit about his belief, claiming to have seen ghosts and citing evidence as to why they exist. In such a case what I possess is a plausible psychological explanation of why, and perhaps also how, that individual came to hold a specific belief, namely due to the role such a belief plays in his psychic economy as a source of hope or comfort relating to his bereavement.

From my point of view, having had such knowledge made available to me, we might think that two things follow, one that is fairly straightforward and one that is more complex. First (a) the specific belief in ghosts, as a knowledge-claim of this individual, would become an object of suspicion and second (b) other beliefs the individual in question holds would also potentially be rendered suspicious. Seeing why (a) is plausible given the above is relatively easy, although also relatively uninteresting, given, as we noted above, that the revealing of flawed methodology in and of itself is not enough to show the belief to be false. However, (b) is more interesting in this context, because if true, then it follows that knowledge of this individual's flawed methods for ascertaining beliefs might make us mistrustful of the individual as a source of reputable knowledge.

Still, seeing why exactly this might be the case is more difficult since the relation between (a) and (b) is not one of strict entailment. The fact that, say, in this particular instance I am suspicious of the reasons for which this individual holds that particular belief (about ghosts existing) need not make me suspicious of all his beliefs. However, what the example shows is that the individual has, at least some of the time, a defective

relation between experience and belief. He is capable of self-deception in the sense of distorting experience (or his reflections on it) so that it can play a role in his psychic economy in a way that makes him, we might say, not properly responsive to reasons. So, our relation to him, *qua* source of reputable knowledge-claims, plausibly mirrors the way in which by having being lied to once by someone I become suspicious and alert to the potential reasons that are being given for other claims they are making; I am, we might say, no longer ‘allowed to believe’ in them (HA 21). In other words, I no longer take at face value the reasons given for the claims that this individual makes and want to know the kind of methodological detail that was revealed to me in the above example. So, mistrust of the individual *qua* source of reputable knowledge seems warranted, given, we might say, I know what he is capable of, or as Nietzsche puts it, ‘the “will” was only too audibly the prompter of the intellect’ (HA 630).²⁶

However, a critic might respond that it would be naïve to think that the beliefs we hold are the product of an entirely disinterested inquiry into the world. Rather, we might think we have certain interests and perspectives, perhaps beginning from some standpoint, and then looking to a certain range of experiential data seek to analyse and test our beliefs, looking for more information, revising our explanations, enriching them with more evidence, and so on. But what we must have, if we are to resist the charge of holding blatantly motivated beliefs in the way the individual in the example does, is openness to the possibility that we might be presented with new evidence that could potentially challenge a belief we hold, and reasonably compel us to a different one. It is this kind of *experiential* method, the one that I argued in Section 1 Nietzsche is committed to, that the individual in the example lacks. Nietzsche himself describes such

²⁶ The claim of being ‘capable’ of self-deception in this context need not imply that the individual is *necessarily* explicitly aware of the way he is distorting experience, but only that he is ‘capable’ of doing so.

a type of person as follows; ‘the presupposition of every believer of every persuasion was that he *could* not be refuted’ (HA 630).²⁷ So, rather than being open to the possibility of his belief being revised by new evidence or by better methods the ‘believer’ is disposed to misrepresent his own experience, perhaps also disavowing certain types of evidence and restricting certain questions in order to sustain the belief regardless.²⁸ These considerations illuminate the idea that revealing a flawed methodology might engender suspicion in the ‘believer’ as a source of reputable knowledge-claims, and it is this kind of suspicion that Nietzsche thinks we should have towards RM.

I said at the end of the previous sub-section that Nietzsche’s critique of RM would primarily target (2*), the amended substantive relation claim. Now, we might wonder how the methodological critique I have framed above relates to RM in this sense. In fact the link is fairly straightforward, since we saw that a methodological critique of this kind reveals a defective relation to experience, such that our individual was no longer a reputable interpreter of experience. And this is what (2*) was aiming at, a certain substantive relation to the world we experience in terms of either an interpretation of experience as underwriting certain metaphysical claims, or supplementary claims about the practical relevance for experience of certain putative metaphysical truths. In this sense *if* RM could be shown to have a flawed methodology in this sense then it would challenge (2*) to the extent that the kind of suspicion Nietzsche wants to engender towards RM would be warranted.

However, even if we accept this argument as valid, Nietzsche still has the burden of

²⁷ See also HA 633.

²⁸ Reginster 2003: 51-85 makes a similar point in trying to understand the *free spirit* as contrasted to the ‘fanatic’. Sartre also says something similar when he comments that the man of *bad faith* ‘apprehends evidence but is resigned in advance to not being fulfilled by this evidence, to not being persuaded and transformed into good faith’ (Sartre 2003: 91).

showing its central premise to be true, namely that RM does *in fact* possess a flawed methodology in his sense, that it is based in ‘passion, error, and self-deception’ (HA 20) in the same way as our individual in above example. Nietzsche in fact expends some considerable effort trying to show just this. What follows are just two of the most striking examples he gives in HA.

First, Nietzsche tells us how those who claim access to the metaphysical world ‘deceive themselves because those things so deeply enrapture them and make them so deeply unhappy’, implicitly presupposing that ‘whatever lies closest to his heart must also be the essence and heart of things (HA 4). Compare this to Schopenhauer’s statement that it is ‘consideration of the suffering and misery of life, that gives the strongest impulse to the philosophical reflection and metaphysical explanations of the world’,²⁹ concluding that ‘if anything in the world is desirable...it is that a ray of light should fall on the obscurity of our existence, and that we should obtain some information about this enigmatic life of ours, in which nothing is clear except its misery and vanity’.³⁰ This, for Nietzsche, is an example of a flawed methodology that is driven by a kind of affect, in Schopenhauer’s case a melancholy about human beings and their lives which, as he concedes, makes the search for specifically a metaphysical explanation, rather than one that limits itself to experience and empirical facts, the driving force of philosophy.

Second, Nietzsche describes how ‘high flying metaphysics’ is suspect insofar as it claims to possess only a pure ‘will to truth’, yet in fact ‘unconsciously’ desires that ‘the significance of knowledge for life *ought* to seem as great as possible’ (HA 6). In this sense Nietzsche claims what is important for RM is not knowledge *per se*, but knowledge that can play a role that is of higher value, namely rendering existence

²⁹ Schopenhauer 1969b: 161.

³⁰ Ibid: 164.

meaningful in the ‘highest’ possible way by giving it a metaphysical value. Presumably then knowledge that would render existence less meaningful, in this ‘higher’ metaphysical sense, is concealed or glossed over. Nietzsche thinks this is evidenced by the way that RM has always had ‘dread for the seemingly insignificant solutions of physics’ and that it plays the role of being an ‘apologist’ for knowledge (HA 6). In a much later work Nietzsche makes a similar point, asking whether ‘it was not really an aesthetic taste that kept humanity in the dark for so long: people demanded a picturesque effect from the truth, they demanded that the knower make a striking impression on their senses’ (A 13). Nevertheless, these passages, and the claims within them, are primarily accusatory, and verifying them might seem difficult.

However, it is telling, for example, that in the Enlightenment of 17th and 18th centuries the vast majority of philosophers, from the Rationalists to the German Idealists, spent considerable time and effort securing a role for some conception of God in their metaphysical systems, in some cases producing sophisticated philosophies of religion.³¹ From Nietzsche’s perspective at least, such projects were clearly motivated by a desire to support certain predominant ideas and as such delimit, prior to philosophical inquiry, what can and cannot be questioned. Perhaps this could serve as a historically verifiable example of RM desire that ‘the significance of knowledge for life *ought* to seem as a great as possible’ (HA 6), insofar as such approaches could be accused of deliberately or self-deceptively obscuring or concealing the import of certain parts of knowledge, namely those pertaining to the development of modern science and historical knowledge. These developments did not, for the most part at least, lead them to question the truth of theistic belief but rather seemed to spur them on to engage in various forms of dialectical intricacy, ostensibly so that the life of human beings might

³¹ See in particular Hegel 1977: 410-478 and Kant 1998b.

maintain a certain metaphysical-religious significance that it had already been decided it ‘ought’ to have. For example, one might consider the way in which despite ‘doubting everything’ in the service of ‘reason’ Descartes still manages to come full circle to argue for a number of, by modern standards, outlandish metaphysical-religious claims (e.g., that we possess ‘immaterial souls’).³² In this sense we might agree with Nietzsche that for RM, maintaining a certain class of beliefs, rather than a genuinely critical or sceptical attitude to knowledge, is important.³³

Perhaps the above examples are enough to bring about that attitude of mistrust towards RM, enough, that is, to discredit RM as a source of reputable knowledge claims and so undermine (2*) the *amended substantive relation claim*. However, even if we grant such mistrust of RM, as far as the canon of metaphysical philosophers prior to Nietzsche goes, could not a new proponent of RM come along and reasonably claim that his methods are not as Nietzsche describes, even conceding, contentiously, that the old metaphysical canon was guilty of the charges Nietzsche levels?

At this point the more positive side of Nietzsche’s critique of metaphysics comes into play. Since, at the moment at which suspicion in RM is engendered, Nietzsche wants to offer us alternative naturalized re-descriptions of the kinds of experiences RM predominantly drew from, that is a re-description of certain important phenomena without any RM postulates. John Cohen captures this aspect of Nietzsche’s critique of RM, telling us that Nietzsche ‘provides an alternative explanatory system at the same time that [he] discredits the metaphysician’s right to defend his or her own’.³⁴

³² Descartes 1996: 24-36.

³³ See also HA 15.

³⁴ See Cohen: 2010: 98. This aspect of Nietzsche’s project shares similarities with that of Hume, for example his dialogues on miracles that have a similar function, i.e., a re-description a familiar source of experience that was claimed on behalf of RM, without recourse to metaphysical explanations (see Hume 1998).

Substantial weight is therefore placed on Nietzsche's ability to successfully provide these kinds of re-descriptions, and Sections 3 and 4 of this chapter will offer central examples of where Nietzsche thinks he can carry out this project, namely with regard to certain moral feelings. So, in sum, at the moment in which suspicion in RM is engendered, in which, in Nietzsche terms, we are 'no longer *allowed* to believe in it' (HA 21), he sets himself the task of explaining how their 'image of the world could differ so much from the disclosed essence of the world' (HA 10).

2.4 Metaphysical Indifference and free spirits

So, given that Nietzsche can provide successful, naturalized, re-descriptions of those experiential phenomena that RM drew from to support (2*), we still need to clarify what is involved in the attitude of metaphysical indifference in relation to TM as the second stage of Nietzsche's critique of metaphysical philosophy. Nietzsche describes the progression to this attitude in terms that fit with the way I have explained his project:

As soon as religion, art, and morality are described in terms of their emergence, so that we can fully explain them to ourselves without taking refuge in the assumption of *metaphysical interventions* at the beginning and in the course of the process, our strongest interest in the purely theoretical problem of the "thing in itself" and "appearance" ceases. (HA 10)

Nietzsche clearly believes the remit of metaphysical philosophy to have been significantly narrowed by his critique of RM, and this seems to have the consequence that purely *a priori* theories, TM, should have no 'strong' hold on us because this kind of inquiry is *not* grounded in experience, which we should, by Nietzsche's lights, no longer grant the metaphysical philosopher licence to draw from or interpret. Given this,

he invites us to conclude that the only type of metaphysical knowledge that we can allow as possible is just not the kind of knowledge that we should be interested in, that is knowledge which is ‘purely theoretical’, where this means it can only be ‘proven’ *a priori*.

Yet, we need to say more about metaphysical indifference in terms of how such an standpoint could be a reasonable evaluative attitude which someone might adhere to, involving, as it seems to, certain judgements of the relative interest and priority of different kinds of knowledge-claims. Moreover, the idea that there are certain class of knowledge claims, those of TM, that we are at liberty to judge as irrelevant might seem contentious or at least in need of further argumentative support. In order to elucidate the potentially problematic nature of metaphysical indifference I will consider an example of a hypothetical conflict between a TM truth and a practical commitment, whereby what is meant by ‘practical’ here is an activity that is dependent on that which is experientially accessible, and so is, in reference to at least some of the beliefs and judgements underwriting it, subject to that which is observable and experientially confirmable. It is also worth bearing in mind throughout the following discussion that, as noted at the beginning of this section, Nietzsche seems to allow for ‘knowledge of a metaphysical world’ as a possibility (HA 9), and the attitude of indifference is reasonably thought to be dependent on there being something contentful to be indifferent towards.³⁵

³⁵ In fact in HA 9 there is an ambiguity in Nietzsche’s formulation about whether or not TM does have content or not, he says that ‘we could assert nothing at all about the metaphysical world except its otherness; it would be a thing with negative characteristics’ (HA 9). If the idea here is that metaphysical indifference follows from the content of TM being purely ‘negative’, then Nietzsche’s attitude might only apply to a limited range of TM, i.e., some kind of neo-Kantian metaphysics which denies our ability to know any positive characteristics about the ‘thing in itself’. However, this is too limited a range of TM to capture the attitude of

I will now set out the example that will guide the discussion. We might ask what the evaluative attitude of metaphysical indifference would suggest if physicalist metaphysics were proven to be true, in that phenomenal, that is experientially accessible, consciousness was shown not to be amongst the most basic existences of the world, such that at the most fundamental level the statement “there is no such thing as consciousness” is correct. At first glance Nietzsche’s recommendation seems straightforward enough. Insofar as someone was indifferent to the truth of any TM, the physicalist claims, even if ‘proven so well’ (HA 9), i.e., proven by purely theoretical or *a priori* means, would not, for his *free spirit*, lead to a substantial revision of practical commitments which seem to depend on the fundamental (e.g., casual) role that phenomenal consciousness plays in them.

A good example of this type of practical commitment might be found in our ideas about the role that conscious thoughts play in reflective deliberation about how to act and related issues of agency (i.e., practical reason). It seems an essential part of human beings self-conception as agents that their conscious thoughts can, at least some of the time, be efficacious in determining how they act, a self-conception we can see quite clearly at work in the important contrast between relatively autonomous persons and schizophrenics who are alienated from the causal origins of their actions (e.g., voices in their head telling them what do). The ability to make this distinction, and the way it is tied to our self-conception as agents, implies that the belief that our reflective

metaphysical indifference (which is surely not just indifference to a specific kind of Kantian scepticism), and certainly would not cover the variety of TM that do involve the positing of some content, which is not merely a ‘negative otherness’, such as physicalist metaphysics (or the positivists metaphysics of Nietzsche’s time). I think that we should interpret Nietzsche’s talk of ‘negative’ characteristics and ‘otherness’ not as a lack of content *per se* but as a lack of predictive content, i.e., empirically verifiable or experientially observable results or consequences.

deliberations can, at least some of time, be efficacious with regard to action, is something we could not be easily persuaded to give up on.

With such an example in mind we can better understand the kind of the conflict at issue and see what is perhaps implied by Nietzsche's remark, in a later note, that 'it is of cardinal importance that one should abolish the *true* world. It is the great inspirer of doubt and devaluator in respect of the world *we are*' (WTP 583), and that 'we do not consider the falsity of a judgment as itself an objection to a judgment' (BGE 4), where the falsity of judgement is to be interpreted as a *metaphysically* false judgement according to a 'proven' truth of TM. This is what Poellner has in mind in his original interpretation of this stance:

Nietzsche's 'free spirits *par excellence*' (GM III 24) are indifferent as to whether the life-world which engages their practical concerns might turn out to be 'fictional' in the light of purely theoretical truths established by a priori metaphysical inquiry. (Poellner 2009: 177)³⁶

³⁶ Peter Strawson suggests something similar when he talks about the evaluative attitude we might take towards the potential truth of determinism in the context of how such a thesis might undermine inter-subjective reactive attitudes like resentment that he takes to play a pivotal role in human life. He says that if we can imagine having a choice between either sustaining practical commitments in some sense recognized as fictional in the light of the truth of determinism, or disabusing ourselves of them entirely, then 'we could choose rationally only in the light of an assessment of the gains and losses to human life, its enrichment or impoverishment; and the truth or falsity of a general thesis of determinism would not bear on the rationality of this choice' (Strawson 2008: 57). For reasons that will become clearer in the discussion that follows I am not sure things are quite so simple. One thing Strawson would owe us is a convincing account of what counts as gains or losses to 'human life', and one might question what other customary practices or habitual evaluative attitudes might be justified on similar grounds.

However, things might be not so straightforward when it comes to this indifferent attitude that Nietzsche's *free spirits* exemplify.³⁷ Since according to the conflict as set out above I seem committed to holding beliefs, and so making judgements, which are not only *prima facie* in conflict, but whose putative truth-status directly undermines the grounds for judgements either implied or explicit in certain other commitments I have, involving me in logical inconsistency. As Poellner states more recently, 'this stance would be incoherent if the [metaphysically indifferent] subjects are construed...as literally accepting a *judgement* as true which they know to be false'.³⁸ To take the example of the physicalist metaphysics again, if I held that it was true, that phenomenal consciousness is 'fictional', it is hard to see how I could make such a judgement and sincerely maintain myself in practices which I believe to depend on its opposite, that phenomenal consciousness is efficacious. Strategies for concealing certain putative truths from myself, say of the physicalist metaphysics, not only seem, by Nietzsche's lights, out of the question, being too similar to the kinds of self-deception he accuses RM of,³⁹ but would generate familiar problems of how exactly such a project would not be self-undermining if it involves bringing to explicit conscious attention the truth that is to be concealed from oneself.

I think we can see more clearly why this problem arises with the first interpretation of metaphysical indifference if we examine the nature of belief more closely and look at the problematic case of deciding to believe. Drawing on a point originally made by Williams, we might think that belief is predominantly a passive state, which, in terms of its propositional content, aims at the truth and as such seeks to, in some sense, represent

³⁷ Poellner concedes as much in a more recent paper (see Poellner 2013: 695-698).

³⁸ Poellner 2013: 698.

³⁹ Although see A 56, 'in the end, it comes down to the *purpose* the lie is supposed to serve'.

reality (belief has, like other cognitive attitudes, a mind-to-world direction of fit).⁴⁰ However, believing *p* is not merely a matter of accepting *p as if true*, since other cognitive attitudes like imagining and hypothesizing also minimally involve accepting their propositional content *as if true*, yet for certain other, i.e., not strictly epistemic, purposes (e.g., hypothesizing might involve accepting a proposition *as if true*, for the purposes of seeing what might follow if a certain claim was believed to be true). Rather, belief more specifically involves, as Velleman puts it, ‘the aim of getting the truth-value of that particular proposition right, by regarding the proposition as true only if it really is’.⁴¹ If this is the right way of understanding belief, as having the constitutive aim of accepting *p* as true only if it is *really is* true,⁴² then it is clear why deciding to believe something, where this involves trying to get oneself to hold a certain belief (say, in the terms of our above example, that the physicalist metaphysics is false, or that phenomenal consciousness is causally efficacious) is an incoherent project.

We can see how by noting that believing *at will* seems to disregard this constitutive aim of belief. As Williams puts it, ‘if I could acquire a belief at will, I could acquire it whether [I thought] it was true or not’,⁴³ and this undermines the status of what it is I have acquired as a belief, as something that aims at representing reality as it *in fact* is. As such, the project would be self-defeating, since both before and after I could not, in full consciousness, hold that what I was seriously aiming at was a belief and that what I then acquired was a belief. Given such considerations I think we should conclude that self-transparent, motivated strategies of trying to getting myself to adopt a belief which

⁴⁰ Williams 1973: 136-151. The idea that belief is primarily a passive state goes back to Hume. For the classical statement of directions of fit see Searle 1983.

⁴¹ Velleman 2000: 252.

⁴² A picture that Nietzsche himself accepts in his statement that ‘every belief is a considering-something-true’ (WTP 15).

⁴³ Williams 1973: 148.

involve what Williams calls ‘truth centred motives’,⁴⁴ whereby what I desire is the *actual* truth of my belief, are incoherent, breaking down under my own conscious attention to them in the sense that I could not legitimately still think I was in the business of believing.

This is not of course to deny that it is possible for individuals to be motivated to believe certain things that they, on *some* level, know to be false, possessing what Williams calls ‘non-truth centred’ motives for deciding to believe.⁴⁵ Yet to do so successfully seems to require either some kind of mental partitioning, or mental gymnastics in order to steer oneself away from countervailing evidence and considerations, since beliefs are, in virtue of aiming at the truth, sensitive in various ways (usually quite strictly) to evidential considerations.⁴⁶ This means that even if I somehow managed to get myself into some state of belief irrespective of its truth, say by drugs, hypnotism or successful self-deception, maintaining the belief becomes almost impossible without yet more mental division and self-deception. As Mark Leon notes, ‘bad belief is like a virus which threatens to poison the rest; either it is susceptible to local correction, that is eliminable in the face of a better-grounded belief or contrary evidence, or its maintenance requires the corruption of the rest’.⁴⁷ Such cases should therefore be thought to represent *global* indifference to the truth of the

⁴⁴ Ibid.

⁴⁵ Ibid.

⁴⁶ One suggestion then might be that insofar as one could have unconscious beliefs that they would, by virtue of not being present to conscious awareness, not be in a position to enter into conflict with those beliefs held at the personal level. However, setting aside whether the idea of unconscious beliefs is philosophically defensible (my sense is that it probably is not), it seems that what would be required is a form of motivated repression of certain specific beliefs in order to escape incoherence and conflict. Whatever metaphysical indifference involves it seems highly implausible that Nietzsche’s *free spirits* are involved in the repression of TM truths (see BGE 39).

⁴⁷ Leon 1992: 312. See also Velleman 2000: 185.

beliefs that are being put forward for assent. Whatever other problems there are with motivated self-deception more generally it seems clear that the project of ‘deciding to believe’ clearly runs the risk of habituated self-division, irrationality and, perhaps if pursued to its extremes, a systematic breakdown of the connection between belief and reality. Clearly, whatever Nietzsche’s *free spirits* exemplify in being committed to metaphysical indifference it should not amount to this.

So if both motivated self-deception and what Williams calls ‘blank projects of belief inducement’ are unsatisfactory,⁴⁸ then what can we say about metaphysical indifference in terms of its stance towards claims that are to be believed to true, like, in our example, the physicalist metaphysics? Framed in terms of the preceding discussion it seems that Nietzsche’s *free spirit* could not be reasonably said to *believe* that the physicalist metaphysics were true and not have this impact on practical commitments dependent on the opposite without such a motivated strategy of self-deception or flouting the requirements of theoretical reason in terms of what constitutes holding a belief.

It is worth pausing at this point to stress that what are usually thought to constitute any knowledge claims, metaphysical or otherwise, are at least three components; justified true belief. We have just seen why the attitude of metaphysical indifference becomes incoherent if it involves a subject entertaining *beliefs* whose propositional content stand in logical conflict (both *p* and *not-p*). As Nietzsche himself notes ‘we are unable to affirm and deny one and the same thing: this is a subjective empirical law’ (WTP 516). In what follows I will consider alternative ways of framing metaphysical indifference that target the different components of a knowledge claim made by a TM; its status as ‘true’ and its status as (rationally) ‘justified’. However, before that it is worth bearing in mind a suggestion that is still focused on the belief component, and is a response to the above problems. It involves denying the status of belief to one side of

⁴⁸ Williams 1973: 151.

the conflict, claiming that some of our practical commitments (those that are in conflict with a certain ‘proven’ TM truth) in fact do not involve beliefs *proper*, but rather have the status of make-believe, such that metaphysical indifference might overcome the above difficulties insofar as it involves a kind of fictionalism.

Bernard Reginster argues for this position. He concedes that whilst, as the above analysis shows, we cannot have beliefs that do not aim at the truth without ceasing to be in the business of believing, Nietzsche’s standpoint towards truth can accommodate instances of what he calls ‘unobjectionable representational attitudes whose content is false or unjustified,’ i.e., of the character of non self-deceptive make-believe.⁴⁹ He claims that Nietzsche is in fact opposed to the practical norm of having one’s cognitive life, that is the representational attitudes present in one’s psychic economy, exhausted or exclusively occupied by propositional attitudes that aim at the truth, i.e., beliefs.⁵⁰ Yet it is unclear if this suggestion will work in the type of case we have been considering, that of a true belief which stands to undermine the grounds for a significant practical commitment.

First, we should note that if what is being suggested is something akin to the mode of imaginative pretence engaged in (mostly) by children, then a constitutive part of successfully maintaining *that* activity involves an awareness that one is, to a large extent, sustaining it by one’s own conscious efforts, on pain of just falling into wholesale illusion or hallucination (something which Nietzsche’s *free spirit* surely wants to avoid). In this sense we might think that convincing pretence, like convincing acting or lying, is, to a least some extent, dependent on knowing the true beliefs that one is flouting and taking deliberate means to avoid ‘breaking character’. In this sense imaginative pretence does not merely involve content we know to be false or

⁴⁹ Reginster 2013b: 448.

⁵⁰ Ibid.

unjustified, say not- p , but involves deliberately representing p when one knows not- p , it includes, as Husserl puts it, a consciousness of a conflict, which no matter how dim, involves one in ‘concealing something in reality’.⁵¹

Yet, the idea that our cognitive representations of reality should involve deliberate make-believe in this sense, and at precisely those moments in which we are confronted with perturbing truths which undermine the grounds of our deeply held practical commitments seems problematic. Could we really live with, that is maintain, practices and commitments dependent on belief in the efficacy of phenomenal consciousness whilst making the judgement “phenomenal conscious is make-believe”? For example, surely we would be very worried about the psychological health of child who was not merely sometimes pretending or engaged in make-believe, but who was doing so at those moments in which there was a challenge to their beliefs and pre-existing representation of reality (say pretending to be a giraffe when they have to attend the funeral of a relative).⁵²

Note though, my claim against Reginster’s suggestion is *not* that it is illegitimate to hold any make-believe representational attitudes whatsoever; the accusation is not of childishness. Rather the claim is that in particularly significant cases in which the grounds for our most important practical commitments have been challenged by a truth, deliberately engaging in make-believe in *those* instances is problematic. Moreover, we might think that pretence and make-believe are typically regarded as harmless attitudes because they come to end; we ‘play’ for a while then to various degrees re-engage in reality. The idea of a person who is habitually, and with regard to issues of fundamental importance for them, sustaining themselves in make-believe practices of this kind seems both psychologically unappealing and clearly runs against the grain of Nietzsche’s

⁵¹ Husserl 2005: 579.

⁵² Reginster fails to note this worry (see Reginster 2013b: 443-448).

claim that ‘the strength of a spirit would be proportionate to how much of the “truth” he could withstand’ (BGE 39; on Reginster’s reading not very much at all). So, fictionalist or make-believe interpretation of metaphysical indifference, which involve essentially abandoning the idea that our practical commitments involve beliefs *proper*, are unsatisfactory.⁵³

Above I said that the attitude of metaphysical indifference might try to target a different component of TM knowledge to explain regarding those truths as irrelevant, or as Nietzsche puts it as the ‘most irrelevant of all knowledge’ (HA 9), and so turn focus away from issues of belief. One potential route then could be targeting the component of truth. In what follows I will explicate this position as having an initial plausibility and then try to explain why it is exegetically and philosophically problematic.

We might think that the language of HA 9, HA 10, WTP 583 and other passages in which Nietzsche expresses the evaluative attitude of metaphysical indifference,⁵⁴ could be read as implying that rather than requiring any explicit conscious effort in order to maintain our practical commitments in the face of conflicting judgments or beliefs, our interest in ‘purely theoretical’ problems ‘ceases’ (HA 10). In this sense it seems as if a certain level of understanding of Nietzsche’s metaphysical critique somehow insulates *free spirits* from such conflicts arising in the first place. Yet we should ask how this could be the case. The suggestion, implied by some of Nietzsche’s comments, is that *free spirits* are to some extent working with a revisionary conception of what constitutes a properly veridical judgement, revealing an underlying commitment to a revisionary conception of what counts as truth *proper*. In this sense they would *not* be accepting a judgement or belief they know to be false in maintaining certain practical commitments

⁵³ In chapter 3 I will also suggest that fictionalist ideas about Nietzsche’s meta-ethics or theory of value are similarly unsatisfactory.

⁵⁴ See also GM III 24, TI ‘How The “True World” Finally Became a Fable’, A 56.

against the grain of certain TM truths, but rather could be seen as having fundamentally revised their epistemological framework in such a way that TM truths no longer count as truths, and so would not be in a position to enter into conflicts with other consciously entertained beliefs.

This seems *prima facie* promising because it avoids the charge of blatant irrationality, since there is no flouting of the law of non-contradiction, which tells us that no statement can be held to be both true and false at the same time; at least in the same sense of true and false. Insofar as the TM truth is in fact not really a truth, or at least not a truth in the same sense as those which underwrite our practical commitments, then the conflict never arises in the first place.

To try to assess the plausibility of this reading we can return again to our example. In line with the above suggestion the statement “physicalist metaphysics is true” would for Nietzsche’s *free spirits*, not count as knowledge insofar as it would not be true in any sense that they are willing to recognize. The limits and boundaries of what counts as knowledge for them being justified *true* belief, where ‘true’ is to be cashed out in terms of being experientially verifiable and so subject only to empirical confirmation and refutation. So, insofar as ‘truths’ established by TM are, as Poellner puts it, ‘propositions which have no predictive consequence and cannot in principle be assessed by our techniques of empirical confirmation or disconfirmation,’⁵⁵ then the suggestion is that they are no longer recognised as potential truths for *free spirits*.

If something like this revised epistemological framework could be defended then it might be thought to provide the kind of insulation or ‘cessation of interest’ (HA 10) that at least some of Nietzsche’s comments on metaphysical indifference gesture towards. So, metaphysical indifference would turn out to be less a kind of pragmatism which might involve weighing up our relative attachment to incompatible beliefs and

⁵⁵ Poellner 2013: 696.

commitments, but rather a revisionary result, dependent on following through the consequences different stages of Nietzsche's critique of metaphysical philosophy.⁵⁶

However, a critic might point out that a revision of our conception of truth in this way essentially reduces metaphysical indifference to positivism, such that at the most basic level statements of the kind "physicalist metaphysics is true", where this is a TM truth, are meaningless. More precisely, it might be claimed they do not make sense insofar as there is no way of verifying or falsifying such statements according to our best methods of empirical confirmation (whatever exactly these are). In other words, whatever such statements amount to they are just not the kinds of things that can have truth-values or a conceivable truth-maker. In this sense metaphysical indifference construed as targeting the 'truth' component of claims to TM knowledge would seem committed to accepting some kind of verificationist epistemology.

Could this really be the heart of what metaphysical indifference amounts to? I think to try to present the position as having at least some exegetical (and perhaps also philosophical) plausibility we should make a distinction between a more methodologically strict kind of logical positivism, and a broader, but also ultimately more vague, experiential-verificationism. The former was typically a strict form of logical empiricism often involving a kind of reductivism that claimed that metaphysical statements were essentially pseudo-statements, and that ethical or aesthetic statements

⁵⁶ This might seem close to Sartre's description of Nietzsche's critique of metaphysics; 'but if we once get away from what Nietzsche called "the illusion of worlds-behind-the scene." And if we no longer believe in the being-behind-the-appearance, then the appearance becomes full positivity; its essence is an "appearing" which is no longer opposed to being but on the contrast is the measure for it' (Sartre 2003: 2, cf. HA 8 and GS 54). However the case of the existential phenomenologists might be problematic, since far from adhering to metaphysical indifference both Sartre and Heidegger claim that phenomenology should, for reasons that are beyond the scope of my discussion here, be a reliable guide to ontology, i.e., to some form of metaphysics (see Sartre 2003: 1-23, Heidegger 2005 1-35).

expressed subjective preferences, neither having any justified place in a philosophically hygienic understanding insofar as they were not verifiable by methods of empirical confirmation enshrined in natural science and mathematics.⁵⁷

As was stressed in considering Nietzsche's methodological commitments in my introduction there are no compelling reasons to think that Nietzsche's philosophy embraces reductivism of this kind. Moreover, this might be seen as a virtue, since the verificationist theory that underwrites positivism of this sort, which aimed to provide a theoretically precise account of the concepts of confirmation and disconfirmation, never managed to do so. The classical statement of some of the central problems are found in Hempel's 'paradoxes of confirmation' which seem to generate counter-intuitive results as to what counts as evidence for a hypothesis. The paradox is generated by showing that from the statement 'all ravens are black' to the contra-position 'everything this is not black is not a raven' we in fact can end up with a statement like 'this red (and thus not black) thing is an apple (and thus not a raven)' as contributing to the body of evidence for the original hypothesis that 'all ravens are black'. As Hempel puts it, we end up implying that 'any non-raven represents confirming evidence for the hypothesis that all ravens are black'.⁵⁸

⁵⁷ For an extended discussion of verificationism see Misak 1995.

⁵⁸ Hempel 1945: 14. Another problem for this kind of empirical verificationism is what existential status to grant to hypotheses about the past for which we cannot refer to empirical particulars (e.g., presently kept documentation). Are we to think that if we do not have the relevant empirical confirmation to hand those events in fact did not happen, or more precisely that we do not have evidence that they happened? As Hempel himself notes neither 'existential clauses' nor specifications of 'fields of application' dispel the paradoxes of confirmation, with regard to the former we might note that scientific hypothesis are proffered in such a way that they can remain purely hypothetical with regard to whether what they hypothesize in fact will 'at some time, actually take place' (ibid 17). Likewise specification of supposed 'fields of application' are problematic, since as Hempel notes 'for a scientific hypothesis to the effect that all *P*'s are *Q*'s, the field of application cannot simply be said to be the class of all *P*'s; for a

Perhaps then we might want to say the kind of verifactionism involved in Nietzsche's metaphysical indifference (that is according to the 'revisionary epistemology' reading) would not be strictly empirical in the above sense but might, in broader terms, *also* appeal to something like descriptive phenomenological claims as contributing to the evidence base for truth-claims. This might seem a promising suggestion insofar as it would connect up neatly with Nietzsche's critique of RM as purging experience of RM explanations (as we saw in Section 2.2, RM wanted to establish that substantive relation to the phenomenal, experientially accessible world). Once so cleansed, descriptive phenomenology could be *part* of a reliable and verifiable guide to potential truth-claims (i.e., 'does it fit the phenomenology'), and, like more strict forms of empirical confirmation for hypotheses, would be open to potential revision and improvement. Such a position would also explain the centrality of self-deception to Nietzsche's critique of metaphysics, insofar as he is implicitly recognizing the experiential realm as a source for potential truth claims, and so in need of considerable cognitive policing.⁵⁹

hypothesis such as that all sodium salts burn yellow finds important applications in tests with negative results; i.e., it may be applied to a substance of which it is not known whether it contains sodium salts, nor whether it burns yellow; and if the flame does not turn yellow, the hypothesis serves to establish the absence of sodium salts'. (ibid 18)

⁵⁹ A number of potential differences are evident between Nietzsche's metaphysical indifference and the phenomenological standpoint of Husserl are worth noting. (1) It is not clear that indifference specifically to TM in Nietzsche's philosophy matches up with the necessary inability, for Husserl, to question the truth-value of our representations (or experiential content generally) when we suspend the 'natural attitude' and are engaged in specifying constitutive conditions for the possibility of an experience. (2) Husserl seems to suggest only a 'temporary' suspension of the 'natural attitude', where the latter involves an interest in the veridicality of our representations, whereas Nietzsche's talk of a cessation of interest or outright indifference seems like something more 'permanent', and quite different from a temporary adoption of a phenomenological standpoint (see Husserl 1970, sections 52-53). The issues here however are suitably complex and beyond the scope of this chapter.

Whilst I think there is something to be said for the above reading it faces exegetical difficulties that render it unsatisfactory. The problem is that no matter what characterization we give of the verificationist principle (either empirical, phenomenological, or some mix of the two) it still involves a significant revision of our sense of truth in such a way that we end up in the position of having to say that the statement “physicalist metaphysics is true” ultimately makes no sense. Put another way, such a reading seems committed to the idea that we cannot really conceive of what it would mean for a statement of a TM like this to be true, and so as to count as knowledge. Yet this seems to contradict Nietzsche’s statement that ‘there could be a metaphysical world; the absolute possibility of it can hardly be contested (HA 9), and that knowledge of ‘such a world’ could be...proven so well’ (HA 9). According to the above reading we would expect Nietzsche to deny the status of knowledge, or even belief, to TM. But Nietzsche does not do this, and so a statement of the kind “physicalist metaphysics is true” (where this is a TM truth) is not, or at least not always, for him, self-contradictory or meaningless, that is to say we can presumably at least conceive what would be required to show this to be the case.

Moreover, as stressed on a number of occasions it seems as if metaphysical indifference would cease to make sense as an evaluative attitude if its object was just nonsensical (the standpoint would then be oddly analogous to being indifferent to the project of trying to imagine a square circle). Put another way, for metaphysical indifference to be a meaningful evaluative attitude worth entertaining requires something *at least conceivable* to entertain as that which it is indifferent towards. If this is right then the above reading, which really just involves ruling out TM *tout court*, is exegetically problematic and potentially also psychologically implausible as an evaluative attitude anyone might be concerned to adhere to.

So, if targeting either the ‘belief’ or the ‘true’ component of TM knowledge seems

problematic perhaps, as a final attempt to render metaphysical indifference defensible, we could target the ‘justification’ component. Generally questions of justification of knowledge, where what we are attacking are the rational grounds for the claims in question, are thought of in terms of a kind of scepticism. In order to see how this sceptical line might work though I think we need to take a step back from the way in which the conflict has been framed up to this point. In a sense what I have been doing is trying to make the situation as difficult as possible for metaphysical indifference, asking what the evaluative attitude could amount to in case a TM was *in fact* proven, such that the statement “physicalist metaphysics is true” is *as a matter of fact the case*. Once set up in this way, we have seen how difficult it becomes for indifference to seem a legitimate stance to take. However, what the sceptical reading might want to say is that whilst it is a *possibility* that something like a TM truth could be ‘proven so well’ (HA 9), i.e., we should not deny, as Nietzsche says the ‘absolute’ (HA 9), i.e., logical and/or conceptual, conceivability of something like this being proven by *a priori* means, that in reality it is very unlikely or improbable that such a proof will be forthcoming.

The sceptic in this instance might say that we have some rather good inductive evidence given to us by over 2000 years of perennial disagreement about nearly all of the potential answers to the fundamental questions in the domain of TM (there never having been anything resembling consensus on what the correct position is in metaphysical philosophy of this kind), for being doubtful that something like a TM could be shown to be rationally justified. Note that what rational justification minimally amounts to here is something approaching widespread consent amongst competent philosophers, who have understood the issues at hand. Something like this kind of response to TM seems to be explicit in Nietzsche’s comment that we should not allow ‘life to hang from the spider threads of such a *possibility*’ (HA 9, my emphasis), and also seems to be part of the logic he presents for this stance in a much later passage

where he remarks, ‘the true world – unattainable? At any rate, *unattained*’ (TI ‘How the ‘True World’ Finally Became a Fable’, my emphasis). This final sceptical reading might fall short of a conclusive blocking of all TM, but given the remit of TM and the way it seeks to establish its claims, perhaps something like the above would be enough to allow for a sensible re-direction of philosophical inquiry away from TM. In other words, given some good inductive evidence about the history and prospects of TM, we might think it reasonable that metaphysics of this kind should no longer be at the forefront of our philosophical inquiries.

However, building on the above, such a sceptical line might also propose that what might be preferable if we want to continue to subject TM to philosophical scrutiny, should neither be the search for better or more logically consistent *a priori* arguments (involving more adroit dialectical maneuvering) against specific TM or more attempted refutations of the possibility of TM. Rather, what might be suggested is something like a moral psychological investigation of what motivates certain TM positions, thereby revealing their claim to ‘cold, pure divinely insouciant dialectic’ (BGE 5) as in fact motivated by something else, allowing one to ‘grasp the historical justification as well as the psychological one for such conceptions’ (HA 20, cf. WTP 583 ‘the “will to truth” [i.e., will to TM truth in this instance] would then have to be investigated psychologically’). This certainly seems to be an important dimension to Nietzsche’s attitude to metaphysical philosophy in both HA and the later works, when he remarks that, ‘to explain how the strangest metaphysical claims of a philosopher really come about, it is always good (and wise) to begin by asking: what morality is it (is he –) getting at’ (BGE 6, cf. BGE 5). Such a ‘genealogical’ approach to TM might involve something similar to the *methodological* critique of RM we considered in the previous section (2.2), and perhaps could engender that broader critical, or in this context we might say sceptical, attitude towards TM we saw in evidence there. But such comments

are primarily suggestive and I do not have the space to explore them in any significant detail here.

In sum, I have tried to explain the two stages of Nietzsche's critique of metaphysical philosophy, the first as a methodological critique of RM, and the second as metaphysical indifference towards TM. While the final position is not without its problems it is clear that something like this stance, as an evaluative ideal, is central to Nietzsche's exemplar of the *free spirit*, involving a liberation from the 'will to metaphysical truth' (GM III 24), and a reorientation of philosophical inquiry to a world which is experientially accessible after it has been purged of RM interpretations (even if, as we have seen, the grounds for this re-orientation might fall short of being rationally required). Nevertheless, and regardless of the problems with metaphysical indifference, the plausibility of Nietzsche's anti-metaphysical stance is dependent on his being able to provide successful naturalized re-descriptions of phenomena that RM appealed to for its claims. The specific phenomena that I want to consider for the rest of this chapter is that of moral feelings.

3. Moral Responsibility and Moral Sensations

3.1 Preliminary Remarks

At the beginning of the second section of HA Nietzsche quotes from Paul Rée's, *The Origin of the Moral Sensations*, stating that 'the moral person, does not stand any nearer to the intelligible [metaphysical] world than the physical person'.⁶⁰ Nietzsche then tells us that that 'this proposition, hardened and sharpened by the hammer blows of historical

⁶⁰ Rée 2003: Section 8.

knowledge, can perhaps someday, at some future time, serve as the axe that gets laid to the root of the “metaphysical need” of human beings’ (HA 37).

This quote introduces the setting for Nietzsche discussion of moral sensations in HA. Nietzsche seeks to dispel the idea that moral sensations might serve as evidence for the substantive relation to experience that RM seeks to establish (2* see above). More specifically Nietzsche has in mind the feeling of moral responsibility (along with the sense of guilt that often comes with this; see Section 4). Nietzsche wants to critique a moral-metaphysical explanation of this sensation by revealing it to be based on a certain, by his lights erroneous, conception of free will. By doing so he hopes to show that we have no rational grounds for feeling a certain type of moral responsibility insofar as it based on an error. Once this error is exposed and we are *also* able to explain these sensations in terms of their historical and psychological role, rather than by appealing to some non-natural conception of free will, Nietzsche thinks we should be able to disabuse ourselves of such feelings (cf. HA 107).

The central passage in which Nietzsche puts forward his account of moral sensations in these terms is HA 39. He tells us that:

The history of those sensations that we use in order to attribute responsibility to someone, that is, of the so-called moral sensations, proceeds through the following primary phases...we make a person successively responsible for his effects, then for his actions, then for his motives, and finally for his being. We finally discover that even this entity cannot be responsible, insofar as it is entirely a necessary consequence, a concretion of the elements and influences of past and present things: hence, that a person cannot be made responsible for anything, neither for his being, nor his motives, nor his actions, nor their effects. With this, we have attained the knowledge that the history of moral sensations is the history of an error, the error of responsibility: as such, it rests upon the error of free will.

(HA 39)

This passage fits with the way Nietzsche's critique of moral sensations was framed above, however it is not entirely clear on what basis Nietzsche reaches his conclusion that 'that the history of moral sensations is the history of an error, the error of responsibility' (HA 39). The bulk of the passage documents the varied ways in which responsibility has been attributed to individuals, from consequentialism ('his effects'), deontology ('his motives') and finally assessment of character ('his being'), yet it is not immediately obvious what critical function this documentation is intended to play even if we grant as plausibly true the empirical premise that different societies, at various times, have attributed responsibility to persons in different ways.

Nietzsche might think that such a history, once acknowledged, stands to undercut the absolute authority of any one of those methods as being the singularly correct way of attributing responsibility. He might also, and relatedly, be expressing scepticism about the possibility of mediating between the different ways of attributing responsibility given that they are ostensibly not based on independent rational argumentation, but are rather the products of specific cultural-historical situations. Yet, even if this is conceded is it hard to see why the 'history of moral sensations is the history of an error' (HA 39), rather than just the history of a variety of ways of attributing responsibility. Moreover, it could be the case that in fact one way of attributing moral responsibility, whilst having a historical origin in some specific cultural-historical situation, has gained independent rational warrant.⁶¹

Given these considerations we might think that reflections on the historical variety of

⁶¹ This problem is different from the 'genetic fallacy' that is sometimes levelled against Nietzsche's critique of morality in terms of the 'origins of morality' in GM. Note that in this passage Nietzsche's aim is not to show a 'disreputable origin' *per se* but rather just to bring to awareness to variety of ways in which moral ascriptions have been made.

ways in which moral responsibility has been attributed will not, or at least not alone, generate the critical conclusion Nietzsche is after; ‘that the history of moral sensations is the history of an error’ (HA 39). Rather, for such a history to play this critical role what is required is the specification of a problematic feature common to attributing moral responsibility in the senses documented. Nietzsche does in fact tell us what this shared component is at the close of the passage, namely ‘the error of free will’ (HA 39). Nevertheless, we might ask what reason he gives us to accept that free will is an error. Moreover, even if we were to accept this, then we would need to be given further argument for the connecting premise, that ascriptions of moral responsibility necessarily depend on free will, such that we cannot imagine or articulate a (perhaps revisionary) conception of moral responsibility that does not depend on free will in the sense that he is opposed to. Nietzsche seems to deny this possibility, saying that ‘a person cannot be made responsible for anything’ (HA 39), a conclusion that seemingly follows, for him, on the basis of ascriptions of responsibility ‘resting’ on the ‘error of free will’ (HA 39, cf. HA 107).

However, such a position might be problematic. After all the *free spirit* seems to attain his exemplary status through a certain conscious effort to overcome certain erroneous assumptions and to adopt certain evaluative attitudes (e.g., metaphysical indifference). Such achievements not only seem to require some conscious effort and understanding but the individual in question might be reasonably thought to be responsible, in some sense, for having undertaken such a project. Clearly the issues are complex enough to warrant clarifying what conception of free will is being rejected here, and relatedly, what conception of moral responsibility.

3.2 Nietzsche's critique of contra-causal-free-will

We can begin to see what conception of free will Nietzsche has in mind if we return to the passage in question. There he tells us, in keeping with his natural-historical methodology (Section 1), that a human being is a 'necessary consequence...a concretion of the elements and influences of past and present things' (HA 39). Whilst Nietzsche does not explicitly say, it is reasonable to assume, given this statement, that when he is talking about the 'error of free will (HA 39) his target is, at least in part, the notion of a contra-causal free will, whereby human beings are thought to possess something like the capacity to be a *causa sui*, that is the ability to be in some sense self-caused.

We can set up the thought as follows. For the will to be free it cannot be caused by anything which is necessarily prior to it, since this would determine it, such that it must be the only and sufficient cause its actions, *ipso facto* a *causa sui*. In a later work Nietzsche frames contra-causal free will, and his opposition to it, in these terms:

The *causa sui* is the best self-contradiction that has ever been conceived...This longing for "freedom of the will" in the superlative metaphysical sense...the longing to bear the entire and ultimate responsibility for your actions yourself and to relieve God, world, ancestors, chance, and society of the burden – all this means nothing less than being that very *cause sui*...(BGE 21, cf. D 124)

So, the reasons Nietzsche provides for rejecting contra-causal free will might be put as follows. He suggests that free will involves the idea of a *causa sui* which is self-contradictory, or as he puts it elsewhere, 'thoroughly absurd' (BGE 15). Presumably this is because what stands to do the causing would not exist to be a cause

prior to some act of causally sufficient self-generation. Moreover, it seems right to say that *nothing* cannot cause *something* to come into existence and insofar as something *is* we must assume something else has caused it to *be* as such.

However, as Hume points out, not having a cause might not in fact be logically self-contradictory or logically inconceivable in this way. Firstly, Hume argues that since we cannot demonstrate with certainty in each instance of a ‘new existence, or new modification of existence’ the productive principle or ‘necessary cause’, then we are not entitled to think that a cause is always necessary.⁶² Secondly he argues that insofar as it is possible in imagination to separate ‘the idea of a cause from that of a beginning of existence’, i.e., the concept of cause from that of effect, such that we can posit the existence or non-existence of an object in imagination without ‘conjoining to it the distinct idea of a cause or productive principle’ then there is no logical contradiction in the notion of a self-cause.⁶³ If Hume is right on these points then Nietzsche’s argument against contra-causal free will as an instance of a *causa sui* might be inappropriate and therefore need to take a different approach.

I think the better way to frame Nietzsche’s argument in this context is by reflecting on what we might think we are committing ourselves to in accepting the notion of a *person* or a ‘will’ that is a *causa sui*, rather than an object of thought, or state of affairs. This more specific focus is captured by Nietzsche’s suggestion that being a *causa sui* would involve ‘pulling yourself by the hair from the swamp of nothingness into existence’ (BGE 21). Needless to say this capacity would involve the power to step outside the natural-causal order insofar as the will as a *causa sui* would not be a ‘necessary consequence’ (HA 39) but would be entirely self-determining and free from all natural necessity. For Nietzsche positing such a capacity is tantamount to

⁶² Hume 1978: 79-80.

⁶³ Ibid.

supernaturalism insofar it would involve denying that human beings are fundamentally natural entities, and that their will (and so their actions), are part of the natural world and occur within it. In this sense Nietzsche thinks that once we accept as plausible a broadly naturalistic picture of human beings, the idea of the will as a *causa sui*, as underwriting a contra-causal conception of free will, will strike us as an implausible ‘superlative metaphysical’ (BGE 21) doctrine, one which, if accepted, commits us to a picture of human beings that is quite radically in conflict with our understanding of human beings as part of the natural world.⁶⁴

However, even if we accept the above there is a second approach for contra-causal conceptions of free will to take that we find in Kantian philosophy. Kant tells us that ‘when we think of ourselves as free, we transfer ourselves into the intelligible world...in thus regarding himself as intelligence man puts himself in another order of things’.⁶⁵ And Schopenhauer says something similar:

Freedom of the will as thing-in-itself by no means extends directly to its phenomenon, not even where this reaches the highest grade of visibility, namely in the rational animal with individual character, in other words, the man. This man is never free, although he is the phenomenon of this will’s free willing. (Schopenhauer 1969a: 288)

What these two accounts share is an attempt to locate free will in a transcendent or

⁶⁴ It is worth nothing that denial of this kind of contra-causal free will has a considerable philosophical pedigree, being put forward in various forms by Hobbes, Hume, Spinoza and Mill, not to mention many others. In fact we find an argument against it similar to Nietzsche’s, dependent on accepting a broadly empirical conception of human being actions and will as part of the natural order, in Schopenhauer (see Schopenhauer 1993: 36). Rée provides a list of those ‘observers of the first rank’ who have denied ‘freedom of the will’ in this sense along with the specific places in which they do so (see Rée 2003: 104).

⁶⁵ Kant 1991: 135.

‘intelligible’ realm, as a kind of freedom that pertains to ‘another order of things’. We can call these variants of contra-causal free will, intelligible free will (in fact HA 39 is titled ‘the fable of intelligible freedom’). Nietzsche frames such intelligible free will as follows:

Schopenhauer believes he can prove the existence of a freedom that people must have somehow possessed, not in respect to their actions, of course, but in respect to their nature: freedom therefore, to *be* this way or that, not to *act* this way or that. From the *esse* [essential nature], the sphere of freedom and responsibility, follows in his opinion the *operari* [work], the sphere of strict causality, necessity, and irresponsibility. (HA 39)

On the Schopenhauerean version at least what such intelligible free will amounts to is pushing back freedom into some notion of character or transcendental self, such that a person’s actions are caused by something that is itself the locus of freedom, namely *free character*.

However, Nietzsche thinks this merely involves pushing the notion of *causa sui* back to the level of character, such that he claims Schopenhauer imagines ‘a person becomes what he *wills* to become, his willing is prior to his existence’ (HA 39). The thought is that someone’s essential nature could only be a warranted object for ascriptions of responsibility for self if the person in question possessed genuinely autonomously control over the nature of that character. Yet, as Janaway notes in considering Schopenhauer’s position, ‘how I could have chosen my own innate, unchanging, non-empirical character by some kind of act lying outside time is never explained’.⁶⁶ The suspicion then is that the only ‘free act’ that could underwrite responsibility for self in this essence-determining sense is something like the capacity of a person’s will to be

⁶⁶ Janaway 2007: 109-110.

causa sui, and that as such all intelligible free will has served to do is push this problematic notion one rung back.

It is worth noting that this criticism applies similarly to non-metaphysically loaded accounts of free will that depend on an analogous idea of freely altering one's character, whereby character then causally determines actions. Since the meta-choices, we might call them, that are made to alter one's character in a certain way as to determine action would presumably also have been caused by something, on pain of being mysterious.⁶⁷ So, such accounts lead to a situation where the only thing that could stop an infinite regress is some meta-act or meta-choice that is a *causa sui*, at which point we are in no better position than the first which claims that it is the will, as the locus of free action, that is a *causa sui*.

Now we have a clearer sense of the kinds of free will that Nietzsche thinks are erroneous we can try to understand what follows in terms of moral responsibility.⁶⁸ Even in BGE 21 Nietzsche talks of the sense in which such conceptions of free will express a desire to 'bear the entire and ultimate responsibility for your actions', at least implicitly holding open the possibility of a partial conception of responsibility that is neither 'entire' nor 'ultimate'. So in order to see what follows from his denial of *contra-causal* and *intelligible* free will we need a better sense of the notion of moral responsibility at work.

⁶⁷ Strawson 1994: 18 makes a similar point.

⁶⁸ Janaway claims, wrongly, that Nietzsche's interest in free will in HA is only concerned with resolving 'the traditional metaphysical debate over the question whether one's actions could have turned out otherwise under identical causal conditions' (Janaway 2009: 63).

3.3 Moral Responsibility as desert responsibility

Ken Gemes provides some important distinctions for thinking about different kinds of responsibility and their related evaluative norms. He tells us:

To say that so and so is responsible for such and such can mean that they deserve punishment/reward for it. On the other hand, to say that someone is responsible for such and such can simply mean that it was their doing. The first kind of responsibility is that which is intrinsically linked to the question of desert, the second kind of responsibility is intrinsically linked to the question of agency. It will be helped then to separate desert responsibility, the kind of responsibility which is a precondition of deserved punishment and reward, and agency responsibility that goes with being the effective agent behind a doing. (Gemes 2009: 34)

Putting agency responsibility aside for the moment, Gemes' distinction seems like a reasonable one, such that at least part of our conception of what it means to hold someone morally responsible for their actions involves a sense of what Gary Watson calls 'reactive entitlement', in that they become ostensibly warranted targets of punishments and rewards.

However, what evidence is there that Nietzsche thinks of moral responsibility as a kind of desert responsibility? In fact, he is explicit that he has such ideas in mind when he talks of moral responsibility, framing his discussion in terms of the concept of 'desert':

Anyone who has completely grasped the doctrine of full irresponsibility can no longer bring so-called punitive and remunerative justice under the concept of justice: presuming that this

consists of giving everyone his due. For the one who is punished does not deserve the punishment...likewise the one who is rewarded does not deserve this reward; he could not have acted otherwise that he did...Neither punishment nor reward is something that comes to anyone as *his due*. (HA 105)

Importantly the final sections of this passage make the connection between this kind of desert responsibility and the kind of contra-causal free will we saw Nietzsche critique. Nietzsche later makes this link even clearer in a passage titled ‘the error of free will’ telling us that ‘people were considered free so that they could be judged and punished’, and earlier in the same passage, that he is ‘describing the psychology that comes into play whenever people are held responsible. - Wherever responsibilities are assigned, an instinct to punish and judge is generally at work’ (TI V 7).⁶⁹

With these passages in mind we can chart Nietzsche’s argument against the notion of desert responsibility as follows. First, we might say that desert responsibility involves adopting a ‘judging’ perspective retrospectively after some act has been committed, whereby ‘desert’ is justified by the idea that someone ‘could have done otherwise’ given the same prior causal conditions. Yet this being able to ‘do otherwise’ given the same causal conditions only makes sense if we assume some kind of contra-causal conception of free will. So the social practice of holding people morally responsible, where this involves a punitive or retributive kind of justice (i.e., desert responsibility) presupposes a conception of contra-causal free will. However, if we accept Nietzsche’s critique of contra-causal free will it follows that desert responsibility, insofar as it presupposes this conception, is rationally suspect. Importantly though, the kind of moral

⁶⁹ See also Z ‘On The Pale Criminal’, and GM I 13. Such judging as a means to punish, perhaps as a way of getting revenge against a party that has caused physical or psychological pain, is one of the key features of the reactive attitude Nietzsche calls *ressentiment* (see GM I 13).

responsibility that Nietzsche's critique of free will undercuts is not responsibility whatsoever.

However, if we abandon such punitive and retributive practices and understand this to be the core of Nietzsche's critique of moral responsibility, such that we should disabuse ourselves of those feelings of 'reactive entitlement',⁷⁰ we might wonder what is left of those evaluative practices of praise and blame. Nietzsche, at times, seems to suggest that his critique of free will and moral responsibility has more dramatic consequences, telling us that someone who accepts the 'complete irresponsibility of the human being' is 'allowed neither to praise nor to blame any longer, for it is absurd to praise and to blame nature and necessity' (HA 107). Yet, it is unclear why this follows, unless rejecting free will depends on the truth of causal determinism, something that was not obviously present in Nietzsche's critique of free will (this might be thought a virtue since many philosophers have thought free will of some sort to be compatible with the truth of causal determinism, e.g., Hume). Moreover, Nietzsche himself retains the practices of praising and blaming, both in HA and across his works,⁷¹ a feature which renders him open to charge of inconsistency if HA 107 is read as referring to praise and blame whatsoever.

So, we need to make a finer distinction between what is involved in praising and blaming, and again we find that the notion of desert responsibility is central to what Nietzsche is opposed to. As Poellner notes:

If what we mean by them [praise and blame] is simply an evaluative judgement – perhaps affectively charged – about an individual in account of their actions and dispositions, then

⁷⁰ Watson 2004: 265.

⁷¹ The most obvious examples might be his negative evaluative judgements of the slaves and the priests in GM I, and his praise of Goethe and Beethoven across the corpus.

there is nothing in principle objectionable about praising and blaming by Nietzsche's lights. If, however, the attitudes referred to by these terms are based on the thought that, among all the causal factors contributing to an action, only those lying within the individual agent are evaluatively relevant, *and* that the agent therefore merits, in addition to the judgements, favourable or unfavourable treatment (reward or punishment) by virtue of her action, then Nietzsche's criticism applies. (Poellner 2009: 151)

However, whilst we might agree with Poellner and Gemes, the latter noting the fact that Nietzsche's tendency to 'praise and blame does not show that he is committed to a positive account of deserts responsibility or deserts free will',⁷² we might then wonder what normative force the evaluative practice of praise and blame retains in the absence of those notions. Susan Wolf highlights this worry in relation to abandoning desert responsibility when she asks how 'persons can deserve a distinctive and more serious kind of blame for being deceitful or petty than pigs deserve for being sloppy or books for being frayed'.⁷³ Yet, this objection might be misplaced if we can articulate a distinctive kind of agency perspective, returning to that agency responsibility Gemes mentions, and a related practice of praising and blaming that does not involve desert responsibility.

First, it might be pertinent to note, as Williams does, that blame, and likewise praise, needs an occasion and an object, typically an action and a person, as a matter of conceptual necessity. Yet, according to Williams, it does not require what he calls the 'pure and isolated form' of a contra-causal free will, which underwrites the demand for retributive justice and which seeks to isolate the person's action from what Williams calls, in terms reminiscent of Nietzsche, 'the network of circumstances in which his

⁷² Gemes 2009: 34-35.

⁷³ Wolf 1993: 64. Cf. Strawson 1994: 5-24.

action was actually embedded'.⁷⁴ As Williams notes, 'the Homeric Greeks blamed people for doing things, and whatever exactly went into their doing so, it was not all this'.⁷⁵ So we need a perspective in which the agent is a genuine presence in his actions, but without the 'isolated' version of free will and deserts responsibility that Nietzsche objects to.

Gary Watson claims that we have such a standpoint in what he calls the 'aretaic perspective' which concerns the question of 'what it is to live a worthwhile life'.⁷⁶ When we adopt this perspective in assessments of agents, we take a view on the kind of life that it being lived out through certain values, projects and ends. Therefore, the kind of judgements we make with these questions in mind will have a different remit and emphasis from those made with contra-causal free will and desert responsibility in mind. They will attend to what Watson calls 'executive and expressive' dimensions of agency,⁷⁷ rather than *causal* accountability, and so need not suppose, given all the same causal antecedents, that the agent could have done otherwise. Nevertheless, our evaluations in these instances will still be about the agents, and their actions, because they will concern the agent's evaluative attitudes and commitments. To respond to Wolf's worry, then, we might say that the praise and blame we express in our evaluative judgements of both ourselves and other agents is distinctive because, in the case of aretaic perspective, it speaks to concerns that we, as human beings, possess and cannot, on pain of living a less meaningful life, fail to take seriously, namely the value of our own lives and the lives of others.⁷⁸

⁷⁴ Williams 1993: 11.

⁷⁵ Williams 1993: 10.

⁷⁶ Watson 2004: 270.

⁷⁷ Ibid.

⁷⁸ I will return to questions concerning the importance of distinguishing evaluative frameworks and self-evaluative frameworks in the next chapter.

Granted that such a perspective is philosophically defensible, and that it can involve praise and blame without falling back into contra-causal free will and desert responsibility, this would be the space that Nietzsche should, in order to avoid the charge of inconsistency, occupy. He could be right that ‘judging is the same as being unjust’ (HA 39) whilst employing a different kind of praise and blame, registering distinctions and making assessments in a way primarily concerned with evaluative practices and commitments. It was these kinds of concerns that I argued in my Introduction are central to his exemplars and the ideals they represent. As even Leiter suggests concerning questions of responsibility in Nietzsche, ‘assuming “responsibility” for oneself is not quite the same thing as actually being responsible for one’s actions. The former is an attitude, a disposition’.⁷⁹ Although we should note the caveat that being responsible for one’s actions can be related to questions of agency, that is whether or not they expressive of one’s agency, and need not be focused on strict *causal accountability* in the sense implied by contra-causal free will and deserts responsibility. Nonetheless, if Nietzschean ‘assuming responsibility’ for oneself is primarily an ‘attitude’ or a ‘disposition’ then this fits neatly with the way I have been framing his philosophy of value thus far, that is, as concerned with expressions of ‘taste’ and their related evaluative commitments, communicated to his readers through exemplars as evaluative ideals.

To conclude this section I want to provide some thoughts about what has been shown thus far and what still needs to be considered. A central consequence of Nietzsche’s critique of free will and moral responsibility is that he believes that, once accepted, it might go *part* of the way to allowing us to liberate ourselves from those ‘moral

⁷⁹ See Leiter 2011: 112. This is something of a qualification to his ‘official’ reading Nietzsche as a hard incompatibilist in the mould Strawson (1994), where this position makes all talk of responsibility essentially meaningless.

sensations' (or affective states) that were tied to these erroneous evaluations, as he notes 'this belief in free will provokes hatred, vengefulness, malice, an entire degradation of the imagination' (HA 99). Aside from those feelings of 'reactive entitlement',⁸⁰ one other affective state closely tied to feelings of moral responsibility is 'consciousness of guilt' (HA 39), and what Nietzsche has not yet provided is a naturalized re-description of *this* feeling which admits of no metaphysical postulates but at the same time can explain its psychological power. This is an important undertaking not least because the 'power' of guilt is often appealed to by moral-metaphysical philosophers as evidence of the need for a metaphysical explanations (i.e., 'guilt as the voice of God in man')

So, if Nietzsche can show us with regard to the feeling of guilt as well that 'the moral person, does not stand any nearer to the intelligible [metaphysical] world than the physical person',⁸¹ then combined with (1) the critique of metaphysical philosophy and (2) the critique of moral responsibility and free will, we will be in a position to disabuse ourselves of the corrupted taste he calls the 'metaphysical need'. To use Nietzsche's own metaphor we would be in a position to lay the axe to the root of the "metaphysical need" (HA 37). So this final section on guilt will complete my account of the *free spirit* and the critical ideal this exemplar represents.

4. Nietzsche's Account of Guilt

4.1 Dissatisfaction-with-self and its development

Nietzsche begins his account of guilt in HA by presenting us with a picture of the way human beings might be thought to evaluate their actions prior to any moralization or

⁸⁰ Watson 2004: 265.

⁸¹ Rée 2003: Section 8.

religiosity. Nietzsche describes how human beings, finding themselves in specific societies and cultures, are conscious of their actions as-to-be-evaluated according to a pre-dominant standard, such that certain actions ‘stand low in the usual rank ordering of actions’ and another ‘class of actions...are generally esteemed as the foremost and highest ones’ (HA 132).⁸² Nietzsche goes on to make the point that for the majority of human beings in such societies there would have been, at least some of the time, a feeling of dissatisfaction when one’s actions did not live up to highest expectations of that society. Nevertheless, he thinks the feeling of dissatisfaction would have only be moderate, and describes it as follows:

This condition would not be felt so bitterly if a person would impartially compare himself only with other human beings: for then he would have no reason to be dissatisfied with himself to any special extent, since he is only bearing the common burden of human dissatisfaction and imperfection. (HA 132)

So, Nietzsche is suggesting that a certain degree of dissatisfaction-with-self is a common feature of socialized human psychology. Moreover this dissatisfaction is only moderate at this stage because when we ‘fall short’ in this way, that is when we fail to perform those ‘esteemed’ actions, we see others around us who do likewise, who are, like ourselves, imperfect. So, according to Nietzsche, there is no ‘consciousness of guilt’ at this stage.

However, something changes when societies move towards what we might think of as a more distinctively moral evaluation of action. Nietzsche describes this stage as

⁸² Nietzsche gives a detailed account of the historical development of ‘bad conscience’ within the ‘social straightjacket’ in the second essay of the GM. The picture presented in HA of individuals in societies evaluating and regulating their actions seems to presuppose something like the ‘internalization of instincts’ Nietzsche describes there.

follows:

How gladly he would experiment with the other class of actions that are generally esteemed as the foremost and the highest ones, how glad he would be to feel himself full of the good consciousness that is supposed to follow from unselfish ways of thinking! But unfortunately it gets no further than this wish: the discontent at not being able to satisfy it gets added to all the other sorts of discontent aroused in him by his destiny generally or by the consequences of those actions termed evil; so that he comes to feel a *profound* discontent...(HA 132, my emphasis)

Here Nietzsche describes a *profound* discontent one feels with regard to one's self, yet it is unclear what exactly this change amounts to. Part of the explanation might come from the idea that rather than a person's actions merely failing to live up to a standard of a certain society they are deemed 'evil', and so as transgressive in a more significant sense. However Nietzsche, in HA, does not give an account of this transition from the evaluative framework of 'bad' to what he considers to be the specifically moral concept of 'evil'. The speculative historical explanations of this transition are not provided, in full, until the *Genealogy*, where the putative 'triumph of slave morality' over 'noble morality' sets the stage for a situation in which a particular set of norms, those of Judeo-Christian morality, became authoritative and are taken to apply absolutely (GM I 1-15).

More broadly though we might say that the moral evaluation of action involves a change in what it takes as its object, whereby what it is assessing is not merely some unsatisfactory action, but rather the whole person. So, the kind of dissatisfaction-with-self we are dealing with then might be thought to have changed from a sense that one merely failed to do something that it would have been 'good to

do', that one fell short in this or any other particular instance, to an evaluation of the entire self as having fallen short. Reginster describes this 'moral' evaluation of action as bringing about the 'ordinary feeling of guilt' whereby 'the diminution of one's worth as a person [is] experienced when one falls short of certain normative expectations'.⁸³ Yet, we need to say something more about this, since we need to understand why, in the non-moral situation, falling short of certain 'normative expectations' did not lead to guilt *qua* diminution of one's worth as a person.

One possible answer might be that within the non-moral framework norm-transgression was seen as more of a happenstance failure of particular actions to match up with certain evaluative standards. In this sense perhaps prior to moral evaluation, norm-transgression was understood to be more just a case of 'bad luck', and so similar in character to Nietzsche's description of a pre-modern attitude to punishment in the *Genealogy*:

He himself, the recipient of punishment, which again descended like a piece of fate, felt no 'inner pain' beyond what he would feel if something unforeseen suddenly happened, a terrible natural disaster, a boulder falling on him and crushing him, where resistance is futile. (GM II 14)

In this sense the dissatisfaction-with-self brought about when I evaluate my action in the non-moral situation is akin to a kind of natural error; something went wrong in that I did not perform the action designated 'good' according to my society, but this does not bring about any *profound* discontent with myself.

We can perhaps make more sense of this when we consider that the moral interpretation of action might involve those concepts of contra-causal free will and

⁸³ Reginster 2011: 57.

desert responsibility discussed in Section 3. Since a more *profound* discontent, as a kind of consciousness of guilt, might emerge when one thinks that, given all the same causal antecedents, one could have done otherwise. In this sense the difference between a non-moral evaluation of action and a distinctively ‘guilty’ interpretation could be understood through Nietzsche’s claim in the *Genealogy of Morality* that ‘the criminal deserves to be punished *because* he could have acted otherwise’, and that this ‘is actually an extremely late and refined form of human judgement and inference, whoever thinks it dates back to the beginning is laying his coarse hands on the psychology of primitive man in the wrong way’ (GM II 4).

Drawing all of the above discussion together then we might put forward something like the following account. With the moral evaluation of an action it is not merely that the action was ‘bad’ and that therefore I am dissatisfied with myself insofar as I have failed in a particular instance to realize a good, but rather insofar as an action is a freely chosen ‘evil’ it reflects strongly on my status *qua* person, whereas in the non-moral stage the relation between the action being ‘bad’, my intention, and an assessment of my worth is much weaker. If this is correct then it seems reasonable to think, as Nietzsche seems to suggest, that the moral evaluation of action would bring about a more tangible feeling of guilt. Yet, it is only with a religious interpretation that we begin to see more clearly how this guilt is intensified, and also what the underlying psychological motivation might be for anyone accepting the moral and religious interpretations of action.

4.2 Religious Guilt and Psychological Motivation

We can become clearer on how Nietzsche thinks the religious interpretation of action works by first setting out how religion relates to questions of dissatisfaction. We might

think that religion responds to that fact about human beings ‘common burden’ (HA 132) as discontented and frustrated with themselves, and then provides an explanation for this (e.g., human beings are necessarily sinful and therefore in need of redemption). According to this account religious solutions are simply a response to something that is already found to be the case. This is how Schopenhauer frames the role of religion, telling us that ‘temples and churches, pagodas and mosques, in all lands and in all ages...testify to the metaphysical need of man, which, strong and ineradicable, follows close upon his physical need’.⁸⁴ In this sense Schopenhauer locates human beings need for metaphysics as a response to the fundamental character of their existence in which ‘nothing is clear but its misery’.⁸⁵

Nietzsche, in contrast, wants to suggest that religion *increases* human beings dissatisfaction with themselves, after which its ‘solutions’ are then offered to a psychology primed for accepting them:

Christianity crushed and shattered human beings completely and sank them as if into slimy depths: then suddenly, in the feeling of complete depravity, the gleam of a divine pity could shine in, so that someone surprised and stunned by grace let out a cry of rapture and for a moment believed that he bore the whole of heaven within him. All the logical excess of feeling, upon the deep corruption of head and heart necessary for it: it wants to destroy, shatter, stun, intoxicate. (HA 114)

In this sense religion would not be a neutral response to an ‘ineradicable’ need as Schopenhauer presents it. Rather, according to Nietzsche, religion is itself responsible for cultivating a psychological propensity to feel more ‘profoundly’ that dissatisfaction

⁸⁴ Schopenhauer 1969b: 361.

⁸⁵ Ibid: 364.

with self we discussed above in such a way that metaphysical consolations then seem to be the only possible solution to a life irredeemably full of suffering. Put simply, religion is, for Nietzsche, as much a cultivator of suffering as it is a response to it.

Nietzsche thinks this psychology of religion is evident in its interpretation of action, which brings us more directly back to the above discussion. He describes the way the religious man evaluates his actions as follows:

He compares himself with a being who alone is capable of those actions that are called unegostical and who lives in continual consciousness of an unselfish way of thinking, with God; it is from gazing into this shining mirror that his own nature appears to him so gloomy, so unusually deformed. (HA 132)

So Nietzsche thinks religion intensifies the moral interpretation of action insofar as it is not just that certain actions are 'evil' and that one should feel guilty for performing them, but that the very possibility of performing 'unegostical' or 'good' actions is withdrawn. So, the ideal of an action which would overcome, at least temporarily, dissatisfaction-with-self is rendered impossible, something we always and necessarily fall short of. As such the individual's sense of guilt is raised to a higher level, since he is guilty for his whole existence which is, by his own efforts, unredeemable. It is at this point, Nietzsche thinks, that religious solutions are offered to a person well placed to accept them.

Yet, we might wonder why an individual would be motivated to accept this interpretation of action. Since on what basis is the divine image, the 'shinning mirror' as Nietzsche puts it, to be understood as something we should be comparing ourselves to, rather than just a religious fantasy? Nietzsche's account here runs the risk of being implausible if it is on basis of the presentation of a divine image that in fact devalues

our existence, and involves a strong feeling of dissatisfaction, that we then have *more* motivation to accept the religion interpretation of action. Why not just reject the devaluation and the increased dissatisfaction-with-self that comes with it in the first place? This point might also apply equally well to the transition from a non-moral to a moral interpretation of action discussed above.

What is missing from Nietzsche's account of guilt in HA is an explanation of what would motivate anyone to accept the intensification of dissatisfaction-with-self from a non-moral to a moral interpretation, and then go even further to a religious interpretation. In the GM Nietzsche provides us with an explanation of why people might have accepted this development in terms of the claim that a certain class of individuals experience an increased feeling of power through it, specifically by engaging in a form of self-directed cruelty:

...that will to torment oneself, that suppressed cruelty of animal man who has been frightened back into himself and given an inner life, incarcerated in the 'state' to be tamed, and has discovered bad conscience so that he can hurt himself, after the *more natural* outlet of this wish to hurt has been blocked, - this man of bad conscience has seized on religious presupposition in order to provide his self-torture with its most horrific hardness and sharpness. Debt towards *God*: this thought becomes an instrument of torture. (GM II 22)

Whilst I cannot do justice to the complexity of Nietzsche's account of 'bad conscience' from the GM, and the way it is intensified into religious guilt, it is clear how the appeal to self-cruelty explains the motivation for accepting the moral, and then the religious, interpretation of action. Since whilst it involves an increase in dissatisfaction-with-self, and so on the face of it more negative affect, this allows for an increase in the feeling of power that according to Nietzsche is experienced when we inflict cruelty. It is only that

in the case of the outwardly powerless, those for whom ‘the more natural outlet...has been blocked’ (GM II 22), the target of cruelty becomes oneself. Janaway captures this thought neatly in terms of his idea of a ‘standing need to express power and hence inflict cruelty’,⁸⁶ which is present in socialised human beings. Yet it is because outward expressions of power through cruelty are ruled out, for the majority at least, that ‘consciousness of guilt’ appeals to human beings in the first place, and therefore can become such a significant psychological means for the satisfaction of this ‘standing need’.

We can further see the distinctive power of the specifically religious interpretation of action when we consider how it becomes a more refined instrument of self-cruelty through the role that a God can play as a ‘judge’ of action. For transgressive actions, so interpreted, are no longer merely contravening against a human state or moral law, but against ‘divinity’. So, for the individual experiencing religious guilt the ‘judge’ is no longer a human one who promises worldly punishment, but is a divine one who, as Nietzsche describes it, ‘hovers before his imagination as punitive justice: in every possible experience, small and large’, promising an ‘immeasurable eternity of punishment’ (HA 132). Religious guilt is therefore, as Reginster notes:

[a] perfect form of self-directed cruelty insofar as it represents guilt as inexpiable. Prior to its Christian reinterpretation, the distinctive feeling of diminished self-esteem expressed by the guilty could only be an imperfect instrument of self-torture because guilt could always in principle be expiated. (Reginster 2001: 77)

So whereas on the moral interpretation of action it was always possible to pay off one’s guilt by suffering some punishment for an ‘evil’ action, in the case of the religious

⁸⁶ Janaway 2007: 147.

interpretation the guilt can never be paid off insofar as man's nature is understood as *essentially* sinful, and so as essentially guilty from the onset. This provides the individual with Reginster's 'perfect form of self-directed cruelty' insofar as it is an endless, or 'inexpiable', source of those feelings of power, always ready to satisfy that 'standing need to express power and hence inflict cruelty',⁸⁷ and as such is, perversely enough, a more desirable interpretation of action than the moral one. It is in this sense that Nietzsche later talks of the individual who possesses a '*will* to find himself guilty and condemned without hope of reprieve, his *will* to think of himself as punished, without the punishment ever being equal to the level of guilt' (GM II 23).

So with this psychological explanation we have made some crucial inroads into understanding 'consciousness of guilt'. Nietzsche highlights the steps, albeit at times opaquely (in HA at least) from non-moral dissatisfaction to moral guilt, and then to religious guilt, through that the idea that in this staged development transgressive actions are given more extreme interpretations; firstly as against state or law, then as against moral absolutes and finally man's very nature becomes an affront to divinity and is interpreted as essentially guilty. Therefore Nietzsche can claim to have provided 'an explanation for the process in the soul of a Christian that we call the need for salvation, an explanation free of mythology: hence a purely psychological one' (HA 132).

Concluding Remarks

The goal of this chapter was to provide an account of the exemplar of the *free spirit* as representing Nietzsche's critical ideal as liberated from the corrupted taste of the 'metaphysical need'. This ideal was characterized initially through Nietzsche's

⁸⁷ Janaway 2007: 147.

two-staged critique of metaphysical philosophy, where it was argued that if his critique of RM succeeds we could progress towards the evaluative attitude of metaphysical indifference with regard to TM. Secondly, we saw how attributions of moral responsibility (as desert responsibility) and the related feelings of reactive entitlement and guilt cultivate the ‘metaphysical need’ both through philosophical errors, such as belief in contra-causal free will, and in perverting and intensifying dissatisfaction-with-self. Having understood both the errors and also the psychological justification for the power of these feelings, the *free spirit* stands to liberate himself from this corrupted taste entirely, representing an ‘innocent human being’ as ‘one conscious of his innocence’ (HA 107). As we saw though this does not amount to the abandoning of responsibility for self whatsoever, but rather just overcoming the desire for desert responsibility and the intensified feelings of guilt associated with the moral and religious interpretations of action.

The hope is that this Chapter has shown how the *free spirit* represents liberation from that affective-evaluative sensibility Nietzsche calls the ‘metaphysical need’ and in doing so opens up the space for Nietzsche’s own positive philosophy of value to be developed. With a view to specifying in more detail what exactly that positive philosophy of value might be I now want to consider the exemplar of Zarathustra and the ethical ideal he represents.

II

The Ethical Ideal: Zarathustra as Exemplar

The aim of this chapter is to give an account of the exemplar of Zarathustra as representing Nietzsche's ethical ideal. Lou Salomé claimed that it is through Zarathustra that 'the highest demands of Nietzsche's ethics find their fulfilment for the first time'.¹ In a similar vein Ridley suggests that it is 'natural to suppose that Nietzsche intended the figure of Zarathustra to be, in some way, exemplary'.² In this chapter it will be argued that Zarathustra exemplifies an ethical ideal in terms of a commitment to self-overcoming (*selbst-überwindung*), although understanding what this involves will require a reconstruction of some of Nietzsche's thinking about first-order, or substantive, ethical commitments.³

To provide some initial setting, some conception of self-overcoming is present in those passages in *Thus Spoke Zarathustra* which stress that 'what is great about human beings is that they are a bridge and not a purpose' (Z P4) and that the 'human is something that must be overcome' (Z 'On Old and New Tablets', 3). Moreover, we might think some notion of self-overcoming is central to both the idea of the overman,⁴ and the recurrent theme of self-sacrifice; 'loving and perishing...these have gone together since the beginning of time. Will to live: that means being willing also for death' (Z 'On Immaculate Perception'); 'I love him who wants to create over and beyond himself and thus perishes' (Z 'On the way of the creator'). Outside this text

¹ Salomé 2001:123.

² Ridley 2007: 101.

³ Zarathustra, being the complex character that he is, might exemplify a number of other things as well as ethical self-overcoming. However, in this chapter I will just focus on self-overcoming, and themes connected to it.

⁴ In using the term overman for the German *Übermensch* I follow the Cambridge translation of *Thus Spoke Zarathustra* (Caro 2006).

Nietzsche also seems to place self-overcoming at the centre of his ethical thought, telling us that his writings ‘speak only’ of his overcoming’s (HA II P1) and that ‘the development of states that are increasingly high, rare, distant, tautly drawn and comprehensive...are dependent on the ‘constant “self-overcoming of man”’ (BGE 257). In later texts he goes on to claim that his humanity consists in ‘constant self-overcoming’ (EH I, 8), and that if someone wanted to give a name to his self-discipline against ‘Wagnerianism’, Schopenhauer, and ‘the whole modern ‘humaneness’ then one might call it self-overcoming (CW P). Finally, he writes that ‘the most spiritual people, being the strongest, find their happiness where other people would find their downfall: in labyrinths, in harshness towards themselves and towards others, in trials; they take pleasure in self-overcoming’ (A 57). Whilst my focus will primarily be on Zarathustra as an exemplar of self-overcoming, insofar as my claim is also that this ethical ideal represents the central positive notion of Nietzsche’s normative ethics I will also draw on other relevant passages.⁵

In terms of structure, Section 1 will give some setting for the importance of self-overcoming by providing a reading of the speeches on the last human being and the overman from the *Prologue* to *Thus Spoke Zarathustra*. Section 2 will then give an account of self-overcoming as a re-evaluative ethical activity, also considering what a commitment to continual self-overcoming might involve. Section 3 will then consider some potential objections to the account.

⁵ For other passages in which Nietzsche praises self-overcoming and highlights its importance see WS 45, GS 357, BGE 224, GM III 16, III 27, TI IX 49, A 38, 53, 59, EH CW 2, EH IV 3, WTP 246, 281, 437, 921.

1 The Speeches of the Prologue

1.1 Preliminary Remarks

By providing a close reading of the speeches of the *Prologue* I will be able to offer some basis for the claim that Zarathustra exemplifies a commitment to self-overcoming, the detailed account of which will then be given in Section 2.

The first thing that we should keep in mind when interpreting these speeches is the setting of the death of God; ‘could it be possible! This old saint in his woods has not yet heard the news that *God is dead*’ (Z P2). I suggested in my Introductory chapter that we should think of the consequences of the death of God in terms of (a) the collapse of any transcendent justification for traditional forms of moral authority, and (b) a shift in the burden of rational warrant onto those who want to sustain the project of traditional prescriptivist ethics.⁶ Therefore in works that come both before and after *Thus Spoke Zarathustra* Nietzsche’s relation to the death of God often takes the form of critiquing certain normative practices which he thinks still implicitly depend on some form of transcendent warrant that can no longer be rationally upheld.⁷

However, in this text his approach is different. This alternative focus is partly indicated when he says ‘here are priests, and though they are my enemies, go quietly past them with sleeping swords...I feel for these priests. And though I also find them distasteful, that is the *least of my concerns* since I have been among human beings’ (Z, ‘On Priests’, my emphasis). Whilst the full meaning of this passage is opaque (after all it does not tell us directly what Zarathustra’s actual concerns are) I think we should see it as suggesting a different emphasis with regard to the death of God. Such that the

⁶ See Introduction Section 3.

⁷ See GS 125, 343.

death of God is indicative of a normative vacuum, whereby the traditional kinds of ethical life are not longer subscribed to. As such, we might say that Nietzsche's central concern in *Thus Spoke Zarathustra* is trying to provide a new kind of normativity, in terms of a positive ideal or ethical standpoint, in a de-deified, post-Christian setting that has taken on board at least some of what Nietzsche himself sees as the consequences of the death of God.

It is this backdrop that informs Zarathustra's speech on the last human being. Since the last human being seems to accept aspects of a de-deified worldview, telling us how 'formerly the whole world was insane' (Z P5). Yet, this figure is presented as the antithesis of the ethical ideal that Zarathustra wants to put forward with his alternative speech on the overman, seen in the way the last human being responds sarcastically to the suggestion for a new ethical ideal, asking 'What is love, What is creation, What is longing' (Z P5). So, in order to understand the new ethical ideal that Zarathustra wants to suggest (and ultimately exemplifies) I will first provide an account of the last human being, which will set up the relevant contrast.

1.2 The last human being

Among the more specific claims made about the last human being are that he will 'no longer have contempt for himself', that he 'lives the longest', 'invented happiness', 'proceeds with caution', and has 'abandoned the regions where it was hard to live' (Z P5). We might think this describes a figure that represents a form of hedonism, whereby the central question of ethical life, of 'what it is to live a worthwhile life',⁸ is answered in terms of a commitment to a set of ends primarily concerned with pleasure, moderation and avoiding suffering; a 'bit of poison once in while; that makes for

⁸ Watson 2004: 270.

pleasant dreams [and] much poison at the end, for a pleasant death' (Z P5). However, labelling the last human being a hedonist might be problematic. Since the above characterization does not fit particularly well with an understanding of hedonism associated with the satisfaction of carnal pleasures, especially if we think of those forms of hedonism which eschew moderation in the search for increasingly intense or excessive states of pleasure.⁹

The last human being in contrast is said to have 'one's little pleasure for the day and one's little pleasure for the night: but one honours health' (Z P5). So we might say that this figure seeks, through prudence and moderation, to avoid the kinds of excess often associated with a more active pursuit of pleasure, that is with a version of pleasure or happiness which is often thought to involve a certain amount of danger, pain and suffering. As Max Scheler notes, 'only in the lowest, most peripheral states of our sensible existence are pain and pleasure found distinct from one other. The *deeper* we probe into ourselves and the more *we concentrate on the actual core of our person*, so much the more do pain and pleasure become less distinguishable'.¹⁰ The last human being would therefore be someone primarily concerned with pleasure as such a 'peripheral state'; 'not your sin – your modesty cries out to high heaven, your stinginess even in sinning cries out to high heaven' (Z P 3).¹¹

⁹ This more active or Dionysian hedonism is exemplified in characters like Wilde's Dorian Grey (see Wilde 2003), and is a preoccupation of Marquis de Sade, particularly his *The 120 Days of Sodom* (Sade 1991). Although this more active hedonism does *not* represent Zarathustra's ideal either (see Z, 'On the Tree on the Mountain'; 'I knew noble people who lost their highest hope. And then they slandered all high hopes...once they thought of becoming heroes: now they are libertines').

¹⁰ Scheler 1974: 134. See also Poellner 1995: 263. Nietzsche sometimes characterizes a similar attitude as a kind of vulgar egoism (see KGW=VII.2.26.262, and BGE 44, 225, WTP 221).

¹¹ It might be tempting to search for a philosophical representative of this sensibility, and Nietzsche's comments about the Utilitarian's are fairly close (see BGE 225, 228) Yet, as Leiter

However, we might wonder why, if all the last human being amounts to is a particularly complacent kind of hedonist, or what Ridley calls ‘the mediocrity of human living with which modernity threatens us’,¹² this sensibility represents such a threat to Zarathustra and whatever his alternative ethical ideal is. Moreover, the speech on the last human being is given with the intention of provoking contempt for this figure, and yet it seems to have the opposite effect, those in the market place asking Zarathustra to turn them ‘into these last human beings’ (Z P5).

To see both the appeal and the problem with the last human being we need to understand how this sensibility represents a particular response to the death of God. We could begin by noting that in a culture devoid of the kind of ethical goals given by externally prescribed Morality and Religion, orientating ethical life around avoiding suffering, abandoning ‘all the regions where it was hard to live’ (Z 5), seems a prudent and rational response. Whilst this is part of the picture we need to understand in more detail how this sensibility relates to that which it is replacing, namely an ethics characterized by moral-religious asceticism. The latter standpoint might be described as holding that the properly moral, and typically Christian, way of life necessarily involves suffering, sometimes to quite a large extent, as shown through the religious martyrs. Yet, one suffers in this life in order to secure a, by atheistic lights imaginary, peace or rest in an afterlife.

However, it is not this goal as such, rest devoid of suffering, or kind of negative eudaimonia, that the last human being objects to. Rather the ‘insanity’ (Z P5) for him lies in the withholding of this state from what is now, after the death of God, the only life there is. In other words, the last human being maintains the moral-religious goal of

comments, such an identification may actually be ‘unfair to *some* Utilitarians (like Mill)’ (Leiter 2015: 105).

¹² Ridley 2007: 90.

negative eudaimonia but naturalizes it, aiming for that passive kind of well-being, absent of all ascetically motivated suffering, in this life.¹³

If this is the right way of understanding the last human being then it is clearer why this figure is problematic for Zarathustra insofar as he wants put forward a new ethical ideal. Since the last human being would be opposed to anything that stands to undermine his primary goal of securing this form of negative eudaimonia in the present. Presumably any ethical ideal that has the character of having not yet been attained, and so representing something that is to be worked towards, presupposes some kind of ascetic dynamic; of suffering now, in some sense, in order to achieve some higher end or value that is not presently secured. Whilst I will return to these points in more detail at various stages in this chapter we can also see that striving towards some higher ideal in this way would involve making qualitative distinctions between certain ends as ‘higher’ and others as ‘lower’, what Charles Taylor calls ‘strong evaluation’.¹⁴ Yet, we might think that these are the kinds of distinctions that the sensibility of the last human being is opposed to, ‘each wants the same, each is the same’ (Z P5), and it is fairly straightforward to see why if what I have said so far is correct. For making such distinctions would more than likely bring about a conflict between his desired-ends, for

¹³ For an interesting account of the relation between ‘well being’ and issues of value see Velleman 2000: 56-84. Whilst I do not have the space to explore the connection between Nietzsche’s and Velleman’s views, it seems the last human being measures the ‘well being’ and the value of his life in terms which Velleman would characterize as primarily additive and reductive. Additive in the sense of primarily being concerned with quantitative considerations, where what is being measured is an accumulation of moments which avoid suffering, reductive in the sense that he is not attuned to the possibility of overarching value judgements about one’s life as a whole. In this latter context Velleman talks of ‘narrative’ considerations which he thinks should not be reduced to the addition of specific moments of eudaimonia.

¹⁴ See Taylor 1989: 14. Remember it was argued in my Introduction that *qualitative* considerations in matters of value are central to understanding Nietzsche’s ideal and his distinction between those affective-evaluative sensibilities (taste’s) he considers ‘corrupted’ or ‘bad’ and those he considers ‘good’ or ‘noble’.

example between an immediate desire for a primarily passive hedonic state and a more variegated desire for some higher ideal, presenting him with potential obstacles or psychological resistances to attaining certain ends, resulting in a sense of frustration, and perhaps feelings of displeasure.

1.3 The overman

Now that we have an understanding of what the last human being represents, we can try to provide an account of the alternative ethical ideal as presented in Zarathustra's speech on the overman. The first thing to note when interpreting Zarathustra's two central speeches on the overman (Z P3, P4) is that they tell us very little of substance about this figure. However, as we will see, this should not necessarily be seen as a problem but rather suggests, in a way that will become more clear in the discussion that follows, that Nietzsche is *not* encouraging people to literally become overmen, in the sense of attaining a determinate state which would instantiate certain fixed characteristics.¹⁵ In this sense I think we should resist Walter Kaufmann's idea that 'the man who has overcome himself has become an overman'.¹⁶ Second, and relatedly, the contrast that is drawn in these speeches is not simply between some overman ideal and its negative counterpart in the last human being, but rather between the latter and those who in some sense take seriously what is implied in the *teaching* of the overman.¹⁷

¹⁵ See Z, 'On the Blessed Isles'.

¹⁶ Kaufmann 1974: 309. Randall Havas also interprets the overman as a determinate future state of human beings who, he claims, have reconciled themselves to certain aspects of temporality. However, the fact that he does not aim to provide a reading of Nietzsche's use of overman in *Thus Spoke Zarathustra* renders his interpretation, and his decision to use the term, whose only significant appearance in the Nietzsche corpus is in this text, problematic (see Havas 2013: 461-484).

¹⁷ Reginster 2006: 251 makes a similar point.

So with these points in mind I will now provide an interpretation of this notion. The central passage reads as follows:

I teach you the overman. Human being is something that must be overcome. What have you done to overcome him? All creatures so far created something beyond themselves; and you want to be the ebb of this great flood and would even rather go back to animals than overcome humans? (Z P3)

The first thing we might note is the obvious contrast with the sensibility of the last human being. There we saw that value was placed on a kind of rest, aiming at preserving human beings in a state of comfortableness. In contrast Zarathustra tells us that ‘what is great about human beings is that they are a bridge and not a purpose’ (Z P4), sounding a similar note to the later Nietzsche’s statement that ‘well-being as you understand it – that is no goal; it looks to us like a end’ (BGE 225), that is, as leading to a cessation of the kind of striving-based ethical life Zarathustra prefers. Clearly, whatever exactly overcoming human beings involves, it would threaten that negative eudaimonia of the last human being.

Building on this contrast I think we can also see that Zarathustra’s teaching of the overman would have to involve making qualitative (rather than merely quantitative) distinctions between higher and lower values, ends and projects, that is, distinctions that the last human being was opposed to making. This is in evidence, first, in the contrast that is drawn between human beings and something higher and more valuable than they are at present, namely some not-yet-present overman. Second, it seems that persons are themselves evaluated by Zarathustra as higher or lower on the basis of their willingness to ‘overcome themselves’, that is to self-consciously will what seems like a piece of speculative anthropology in the idea that ‘all creatures so far created something beyond themselves’ (Z P3). This is clearly an esteemed project in the light of which Zarathustra

makes negative assessments of those who fail to take it seriously, criticizing those who represent the ‘ebb of this great flood’ and would rather ‘go back to animals’, that is would prefer a life characterized by that passive kind of well-being of the last human being, rather than one lived in the pursuit of ‘higher’ ends.

However, it is still relatively unclear what the ‘higher end’ of the teaching of the overman really involves, and what the figure of the overman more concretely represents. In order to try to get to grips with this I want to first provide some reflections on that fact that the teaching of the overman involves something that is often framed by Nietzsche in terms of a rhetoric of sacrifice. We find with this idea in Zarathustra’s descriptions of ‘going under’ and ‘perishing’:

I love those who do not first seek behind the stars for a reason to go under and be a sacrifice, who instead sacrifice themselves for the earth, so that the earth may one day become the overman’s...I love all those who are like heavy drops falling individually from the dark cloud that hangs over humanity: they herald the coming of the lightning, and as heralds they perish. Behold, I am a herald of the lightning and a heavy drop from the cloud: but this lightning is called overman. (Z P4)

What should be noted here it is that the theme of sacrifice implies a certain ascetic dynamic, indeed in a later passage Nietzsche even writes that those who take pleasure in self-overcoming are ascetic by ‘nature, requirement, and instinct’ (A 57). This has led certain commentators to claim a close resemblance between Zarathustra’s teaching and religious asceticism, even going as far as to claim that Nietzsche’s concept of self-overcoming is a Christian notion.¹⁸

¹⁸ Kaufmann agrees with Klages on this point, telling us that ‘Nietzsche...would measure the value of conduct in terms of self-overcoming: this *Überwindungsmotiv* is, as Klages insists correctly, Christian’ (Kaufmann 1974: 260-261 and Klages 1926).

Yet, this interpretation is mistaken. Since, whereas in previous religious asceticism sacrifice was directed towards, and rendered meaningful in reference to, securing for oneself some final state of eudaimonia (the end which I argued is naturalized by the last human being) the teaching of the overman suggests a different kind of ‘this-worldly’ asceticism, what Zarathustra describes as a ‘sacrifice...for the earth’ rather than a sacrifice whose reasons are ‘behind the stars’ (Z P 4), i.e., an asceticism motivated by a religiously transcendent ideal. This contrast is made clearer if we note that the kind of asceticism present in the teaching of the overman in fact opposes the Christian notion that the individual is sacrificing himself for some transcendent future state, at which point he stands to attain some kind of peace or rest. Rather, the ‘sacrifice for the overman’ is made in such a way that the one who makes it ‘wants no thanks and gives none back: for he always gives and does not want to preserve himself’ (Z P4). So, whatever asceticism is present in the teaching of the overman, it is not, to use Nietzsche’s later formulation, ‘life against life’ in the service of some ‘quite different kind of existence that it opposes and excludes’ (GM III 11, 13), since it opposes any sacrifice for a final state of desireless rest, metaphysical or otherwise, that is the end of both religious asceticism and the last human being.

So, we might say that Zarathustra’s teaching radicalizes the dynamic of asceticism by withdrawing any once and for all satisfaction, representing the highest end of ethical life in a temporally distant and personally dislocated way, that is, in the sacrificial preparation for some future overman. Therefore the value of an individual human life, where this might take the form of a secured meaning either in some satisfactory present or imagined future, is removed. It is unsurprising then that Zarathustra describes his teaching of the overman as involving a kind of ‘madness’ (Z P5). However, behind this rhetoric, these reflections point towards a more subtle understanding of this figure. We might say, prefiguring more discussion below, that the overman in fact represents for

Nietzsche the ‘horizontal’ nature of ethical life, where this involves a commitment to some kind of *continual* self-overcoming in pursuit of our highest ideals. Reginster recognizes this point, although framing it in the terms of his understanding of self-overcoming as a kind of power-ethics:

Nietzsche never offers a substantive characterization of the overman, because it represents the indeterminate, ever-receding formal objective of the individual who, by virtue of engaging in the pursuit of power, is perpetually in search of new challenges to meet, of new overcomings. (Reginster 2006: 251)

At the start of Section 2 I will suggest that we can provide a different interpretation of what the commitment to self-overcoming involves that is not exclusively characterized by the pursuit of power in terms of overcoming resistances. However, for the moment I think we can see that Reginster is right to think that (a) Zarathustra’s teaching of the overman should be interpreted as suggesting that the highest commitment of ethical life should be to some kind of self-overcoming and (b) that there is something in this notion which clearly opposes any once and for all satisfaction, presenting the highest ends of ethical life as somehow ‘ever-receding’ or what I will call horizontal.

I will provide more substance to these ideas in Section 2, however even framed in these relatively sparse terms this interpretation allows us to see how some of the more standard objections to Zarathustra’s teaching of the overman might be misplaced. Maudemarie Clark, for example, claims that Zarathustra’s teaching ‘appears to value human life only a means to a superhuman life’ concluding that ‘the overman ideal expresses Zarathustra’s own need for revenge’,¹⁹ and Karl Jaspers, that it involves a ‘biological conception of a method of breeding and cultivation from which we could

¹⁹ Clark 1990: 273-275.

expect that a new being would arise at the boundary between the existing species of man and a higher one'.²⁰ However, both these objections misunderstand that the concept of the overman does *not* represent some determinate future biological object, but rather represents a commitment to self-overcoming and, more subtly as I will try to explain in later sections, metaphorically represents something important about the way in which Nietzsche thinks we strive towards ethical ideals (i.e., some sense of never being quite satisfied with what we are at present and always falling short of that which we were striving for; see Section 2.4).

Yet, Clark and Jaspers are responding to something that is present in Zarathustra's speeches, namely a certain biological, quasi-evolutionary rhetoric. Given some of what is said there it would be easy to conclude that we are being impelled to either make of ourselves or prepare the way for some, as Poellner critically observes, 'really possible future entity, some special kind of post or super-human biological object'.²¹ I will try to explain below why the language of the *Prologue* sometimes takes this form (see 1.4), however, in line with the above interpretation, I think we do best to understand the teaching of the overman and the overman rhetoric in general to refer to, as Poellner puts it, 'aspects of humanity itself which self-consciously exceed any such self-objectifying categories'.²²

Whilst I will return to the philosophical detail of these points in later sections, it is worth keeping in mind that a crucial reason why Zarathustra frames his ethical ideal of self-overcoming through the teaching of the overman is because of this device's ability to highlight the way in which the nature of self-overcoming, as a transformative

²⁰ Jaspers 1997: 168. Although Jaspers position seems confused between this formulation, and one which seems closer to the interpretation I want to argue for, where he says that 'the image of the superman, as Nietzsche sees it, remains indeterminate. The weight of his thought lies in the task which he assigns' (Jaspers 1997: 168).

²¹ Poellner 2012: 79.

²² Ibid.

activity, never allows us to be done with ourselves. In this way by framing his ethical ideal through the overman Nietzsche reveals his resistance to any construal of self-overcoming, and ethical pursuits more broadly, that would locate their ultimate value in some achieved, ‘self-objectifying’, end state, that could be definitively achieved. In this sense then, the importance of teaching of the overman is not to be located merely in its opposition to the complacent eudaimonia of the last human being, but also, it will be argued, suggests something interesting about attempts to realize our ‘highest’ ethical ideals. Nevertheless, these are complex points, and to make more sense of what they might imply I will first have to provide a more detailed account of self-overcoming. However, we can see why the above reflections imply that ‘biological, sociological, or psychological concepts’²³ such are those expressed in Clark’s and Jaspers’ interpretation of the overman would most likely betray Nietzsche’s intention.

1.4 The Prologue and the move to an exemplary role

Before providing a detailed account of self-overcoming I want to offer some reflections on the provisional nature of Zarathustra’s speeches in the *Prologue*. After giving both speeches, which fail to rouse those in the market place, Zarathustra is shown by Nietzsche to draw some important conclusions about why his speeches miscarried. The central point is that Zarathustra mis-selected his audience, saying after the speeches, ‘they do not understand me. I am not the mouth for these ears’ (Z P5). This is then explained in a longer passage in which Zarathustra is said to have discovered ‘a new truth’:

²³ Ibid.

It dawned on me: let Zarathustra speak not to the people, but instead to companions! Zarathustra should not become shepherd and dog of a herd...I do not want to even speak again with the people...I shall join the creators, the harvesters, the celebrators: I shall show them the rainbow and all the steps to the overman. (Z P9)

So, Zarathustra originally spoke to ‘the people’, and he concludes that this was a mistake, making a resolve to abandon this kind of preaching *en masse*. The rationale for this seems to be that there is something about ‘the people’, that is the communal market place of ideas, which in a post-Christian age, is resistant to the kinds of things being suggested by Zarathustra. In this sense he, at first, might be said to misunderstand the problem that humanity faces, as if the issue is merely the lack of a new ideal, or a kind of ethical lethargy, in the wake of the death of God. As Pippin writes:

[what humanity suffers from is not] merely a fearful failure of will, a failing that requires a rhetoric of courage, a call to a new kind of strength...the problem Zarathustra confronts seems to be a failure of desire; nobody wants what he is offering, and they seem to want very little other than a rather bovine type of happiness. It is that sort of failure that proves particularly difficult to address. (Pippin 2006b: xxii)

So, this confusion on Zarathustra’s part about what exactly his teaching was a response to, and whom it was a response for, obscured the import of what was said. We can see this in the fact that because Zarathustra was speaking to the ‘people’ his teaching was framed in an overly emancipatory tone, sometimes lapsing into that literalist rhetoric of speaking as if the overman was some future biological objective for present human beings to realize.

So, as Nietzsche makes clear, the failure of the speeches in the *Prologue* serves to bring to awareness the problem with Zarathustra occupying the traditional role as a

moral pedagogue and of the audience as the passive receptacles, who might take up his new teaching. Zarathustra's failure to achieve his initial aims therefore allows him to draw significant conclusions about the way in which his project should continue, whereby he is shown to have decided against the 'shepherd' role in favor of a more exemplary one, that of joining 'the creators, the harvesters, the celebrators: I shall show them the rainbow and all the steps to the overman' (Z P9). In this manner it makes sense, given some of the reflections about the important role of exemplars in Nietzsche's philosophy of value from my Introduction, that Zarathustra should be seeking 'companions who follow me because they want to follow themselves...I want to go to my goal, and I go my own way' (Z P9). Moreover, this move to a more exemplary role is also significant because it allows for a change from Zarathustra as merely the teacher of the overman, to something more interesting that then takes shape in the rest of the text, namely Zarathustra as an exemplar who represents a commitment to the ethical ideal of self-overcoming.

2. Nietzschean Self-overcoming

The aim of this section is to provide an account of what is involved in self-overcoming. Firstly, I will consider some reasons for resisting framing this concept exclusively in terms of will to power psychology. Secondly, I will present my own account of self-overcoming as a distinctive kind of re-evaluative, ethical activity. Thirdly, I will give an example of this kind of self-overcoming through Nietzsche's idea of achieving a perspective 'Beyond Good and Evil'. Finally I will consider the idea that we should be committed to self-overcoming continually, and what this aspect of Zarathustra's ethical ideal might amount to.

2.1 Will to power and Self-overcoming

A number of recent commentators have explained self-overcoming, and this concept's relation to Nietzsche's ethical ideal, through psychological versions of the will to power. On one prominent version, as put forward by Reginster, we are given an account of human desire in which the ends of our first-order desires are sought in conjunction with a second-order desire for the feeling of power; a hedonic experience occasioned by overcoming resistances.²⁴ Whilst this second-order desire, the will to power as the will to overcoming resistance, is dependent on those first order desires, insofar as it can only get determinate content and occasion for expression through them, this psychological will to power has been put forward as fundamental to understanding Nietzsche's ethical ideal and philosophy of value. According to these interpretations, the kinds of first-order activities and ends that Nietzsche thinks are most valuable, and wants to direct us towards, are those in which the will to power so construed is maximized.²⁵ In this way it is claimed that will to power psychology can provide us with a standard by which we can assess and rank human activities, underwriting an 'ethics of power'.²⁶

For these commentators self-overcoming is framed in terms of this will to power

²⁴ Reginster 2006: 118-147. For passages that support this reading see WTP 688, 689 and A 2.

²⁵ Aside from Reginster, Katsafanas also provides a similar account, although with more agential and 'constitutive' themes in mind (See Katsafanas 2011: 650-651; Katsafanas 2013a 145-183). Richardson, with a slightly different emphasis, states that, for Nietzsche, 'I am to see that the point to life is growth, growth by overcoming previous states of myself...Nietzsche's view again looks like a kind of consequentialism, with power – the individuals own power – as the good to be maximized' (Richardson 2013: 777).

²⁶ Reginster 2006: 176. It should be noted that for Katsafanas will to power psychology only gives us a standard according to which we assess *other* values, and with which these values need be consistent (i.e. not in conflict), rather than a foundational principle from which we might derive all other normative claims (see Katsafanas 2013a: 189).

psychology,²⁷ such that a commitment to self-overcoming is a commitment to perpetually maintaining this power-dynamic of seeking out and overcoming resistance in the suitable first-order activities. Reginster, for example, tells us that Nietzsche's creators will power in this way, they are 'perpetually in search of new challenges to meet, of new overcomings', and in doing so eventually overcome themselves.²⁸ However, Reginster does not say enough about what is involved in this process. Both his and Paul Katsafanas's official position seems to be that self-overcoming is another term for the will to power, and as such is (like will to power is for them) a constituent feature of all human psychology. Yet, at other times they, like Nietzsche, describe self-overcoming as a distinctive achievement, and so we are left with a number of questions. One might wonder if it is the case that the achievement of self-overcoming is merely quantitative, such that in overcoming enough resistance one overcomes oneself, or whether there could be degrees of self-overcoming. Moreover in such instances when 'creators' eventually overcome themselves we might ask whether this was an explicit aim of the activity or merely a by-product.

In fact, the way Nietzsche sometimes talks about self-overcoming implies that it is not sufficiently explained as resulting from the pursuit of overcoming maximum resistance. As highlighted by the quoted passages from both *Thus Spoke Zarathustra* and Nietzsche's other works at the beginning of this chapter we might think that it also matters to Nietzsche what it is that stands to be overcome. As Poellner notes, for Nietzsche 'evaluative judgements are also possible and required about the worth of an overcoming in terms of what it is an overcoming of and what it is an overcoming

²⁷ Katsafanas 2013a: 158; Reginster 2006: 250-251; Richardson 2013: 777. This reading of self-overcoming was originally proposed by Kaufmann, who writes that 'the will to power is conceived of as the will to overcome oneself', and that "'Power" means something specific for Nietzsche: self-overcoming' (Kaufmann 1974: 200, 261).

²⁸ Reginster 2006: 250-251.

towards; that is, what our first-order desires aim at'.²⁹ That is to say we might think that Nietzsche's evaluative commitments, whilst often plausibly framed in terms of will to power psychology and its conception of self-overcoming, are *also* concerned with first-order evaluative practices in terms of their substantial content and normative ends, what in my Introduction I called the qualitative aspects of Nietzsche's ideals.³⁰

In the following section I will argue that by interpreting self-overcoming as a distinctive kind of re-evaluative activity we can provide a different account of this concept and its role in Nietzsche's ethics than in those accounts that focus predominately on will to power psychology. However, it should be kept in mind that I am *not* claiming that the accounts which interpret self-overcoming in that way are wrong to do so, or that my reading should replace those accounts. As both Reginster and Katsafanas argue, there are good textual and philosophical grounds for their reading. In this sense what follows of my own account of this concept is *not* intended to be a sufficient characterization, either exegetically or philosophically, and in this sense I do not claim to capture everything, or even necessarily the most important thing, that Nietzsche intends with this concept. My claim is more limited. Namely that my reading of self-overcoming as a distinctive kind of re-evaluative ethical activity points towards something that the will to power accounts potentially overlook and therefore warrants consideration.

²⁹ Poellner 2012: 55.

³⁰ As Brian Leiter comments, one of the central charges that Nietzsche makes against the kinds of morality he rejects are their 'normative agenda[s] which benefits the "lowest" human beings while harming the "highest"' (Leiter 1997: 263). For a contrasting approach that makes 'impotent' will to power central to Nietzsche's critique of morality see Reginster 2013a: 701-727.

2.2. *Self-Overcoming as a re-evaluative activity*

So, we need an account of self-overcoming that says something about ‘the worth of an overcoming in terms of what it is an overcoming of and what it is an overcoming towards’.³¹ In other words, we need an account of the way self-overcoming relates to questions of value both in terms of (a) its substantive content and (b) its ends. In this sub-section, I will provide the start of an account of (a). The central claim I want to argue for is that self-overcoming involves a distinctive overcoming of values in relation to what I will call *self-evaluative* frameworks. What I mean by this will require explanation. To begin with it will be helpful to have some idea of what constitutes an evaluative framework *simpliciter*.

First, evaluative frameworks should be thought to involve normative standards that allow for judgements of actions (or states of affairs) that fall under the purview of the activity in question. Games are good examples of when we adopt such evaluative frameworks, since by having in place certain pre-established normative standards about what the relevant ends are, and how we are to achieve them, they allow for contrasting, and often finely grained, judgements of performance. A second feature of evaluative frameworks is that insofar as we are sincerely engaged in activities governed by them, then the normative standards they provide are *constitutive* standards, such that if I do not follow the constitutive standards then I cease to performing that activity. For example, if when playing cricket I claim to have scored a run by jogging on the spot, this indicates that I am not really playing cricket at all.

With this characterization of evaluative frameworks in place, I will try to specify what the relevant differences are in the case of *self-evaluative* frameworks. First, it should be noted that the distinction cannot just be that the activity in question involves

³¹ Poellner 2012: 55.

self-evaluation, such that it is specifically my, or someone else's, performance that is being assessed, rather than that of an object (e.g., a car engine's performance). Since, evaluating people according to externally specified normative standards is typical of, and essential to, the vast majority of human activities.³² Rather, the important distinction to be drawn is to do with the content of what is being assessed and the type of questions this different focus generates, such that we might say that self-evaluative frameworks concern others and myself in an ethical sense. Taylor's description of 'strong evaluation' sets up some criteria for distinguishing self-evaluative from evaluative frameworks on such grounds:

There are questions about how I am going to live my life which touch on the issues of what kind of life is worth living, or what kind of life would fulfill the promise implicit in my particular talents, or the demands incumbent on someone with my endowment, or of what constitutes a rich, meaningful life – as against one concerned with secondary matters or trivia. (Taylor 1989: 14)³³

So, we might argue that with self-evaluative frameworks we are under the jurisdiction of a set of normative standards according to which assessments are made not just our

³² Leiter makes this error when trying to define what constitutes an instance of a morality for Nietzsche. He appeals to the idea of 'Anthropocentric Evaluative Practice', the practice of evaluating oneself and others, claiming that 'both slave and master moralities are examples of *morality*' because they are evaluative practices that are 'concerned not with things or texts or foods, but with *human beings*' (Leiter 2015: 138) Yet, this does not tell us what is distinctive about moralities or (in my terms) self-evaluative frameworks. So whilst he rightly adds that '*not every AEP is a morality*', it cannot therefore follow that slave and master moralities are examples of morality merely *because* they involve AEP. Rather AEP would be a necessary condition, but is not sufficient. As such Leiter's definition does not tell us what distinguishes self-evaluative frameworks (his moralities) from activities that are AEP's, such as games. Gemes and Janaway also highlight this problem (see Gemes and Janaway 2005: 737).

³³ See also Frankfurt 1998: 92 and Watson 2004: 270.

performance in any particular activity, but of our conduct in ethical life. In this sense the normative standards that constitute self-evaluative frameworks will be more general, but also more fundamental and important. For what self-evaluative frameworks are comprised of are the criteria according to which we make assessments of our worth as persons, what we might call our most fundamental evaluations, and therefore the normative standards through which we construct our practical identities.³⁴ As Nietzsche himself writes in a note, ‘morality is the doctrine of the order of men’s rank, and consequently also of the significance of their actions and works for this order of rank: thus, the doctrine of human valuations in respect of everything human’ (WLN 35[5]).

Another significant aspect of self-evaluative frameworks are normative ideals against which we compare ourselves, often as a means for self-improvement.³⁵ Consider, for example, the self-evaluative framework of the Christian. This individual’s assessments of ethical conduct are bound to an ideal Jesus was taken to exemplify, providing the Christian with certain fundamental evaluations. I note this feature because we might think that a self-evaluative framework’s normative ideal provides at least some of the content of those, often contrastive, fundamental evaluations. We can see this in the case of Christian Morality through its evaluations of redeemed vs. condemned, good vs. evil. Yet, whilst such normative ideals for self-improvement are a central feature of self-evaluative frameworks it is what they concern that is important, namely, as stated above, assessments of ethical conduct in a way that is bound up with our practical identity.

³⁴ I will explain my use of this latter term and its importance for my account of Nietzschean self-overcoming in what follows.

³⁵ This feature is present in some evaluative frameworks but not in others. For example, think of evaluative frameworks for assessing literature. Since whilst we might have some ideal of what the highest type of literature should attain to, in assessing a particular work we are *not* always doing so as a means to improve it.

Pausing for a moment, we might wonder what to make an individual who might value his performance in, say, a game so highly that it defines his practical identity. Such an individual would not obviously be shallow in Taylor's sense of failing to be attuned to issues of what it is worthwhile to do and what is worthwhile to be, since they would have answers to these questions.³⁶ However, we might think that one of the central intuitions which guides our sense that self-evaluative frameworks are more important is that they should be comprised of a more general set of fundamental evaluations that stand to determine our attitude towards a wide variety of more specific first-order projects or ends, so determining 'the significance of their actions and works for this order of rank' (WLN 35[5]). Therefore, it might seem problematically narrow if any one particular activity exclusively determines these fundamental evaluations.³⁷ So there is an important issue of the scope of one's fundamental evaluations, such that if, as in the above example, an individual is *only* concerned with his performance in a particular game at the expense of all other concerns then we might think it an open question whether or not he is really engaging in 'ethical life' as we understand it.³⁸

Now that we have a working definition of a self-evaluative framework I can return to my initial claim. I want to argue that self-overcoming, as an ethical ideal, involves both questioning and ultimately abandoning a self-evaluative framework, that is overcoming some set of fundamental evaluations through which we make assessments of our ethical

³⁶ Taylor 1985: 26.

³⁷ See Taylor 1985: 39.

³⁸ Perhaps this will seem a question begging response, or a merely stipulative definition (i.e., defining this individual 'out' of ethical life), to someone who is willing to claim quite baldly, that concerns relating to his performance in the game do in fact provide him with a wide range of evaluative attitudes towards other first-order projects and ends. For example he might be willing to entirely disregard certain other-directed concerns in favour of focusing exclusively on training to improve his performance. In the end if such an individual does not strike us as something of an un-ethical sociopath I expect that we cannot do much more than appeal to our intuitions about the necessary scope of ethical life.

conduct, therefore shifting the normative standards through which we define our practical identities. It seems that Zarathustra has something like this in mind when he says that the ‘greatest thing that you can experience’ is ‘the hour of your greatest contempt’ (Z P3), suggesting the creative power of certain kind of dissatisfaction-with-self.³⁹ So, we might say, framed in the terms of the first desideratum set out at the beginning of this section, that self-overcoming takes self-evaluative frameworks as its substantive content, that is what it is an overcoming of, and as such it is a re-evaluative, ethical project.⁴⁰

The exegetical case for attributing this view of self-overcoming to Nietzsche, while to my mind present in many of the passages quoted in the introductory section of this chapter, will be provided in more detail through a central example in Section 2.3. However, before this I want to provide some further reflections on two aspects of my account thus far. Firstly I want to discuss what kind of activities this interpretation of Nietzschean self-overcoming rules out as counting as an instance of this achievement, and secondly how the account presented so far makes sense of the ‘self’ in self-overcoming.

It is important to note that interpreting self-overcoming in my way rules out many first-order activities, and their ends, as counting as instances of Nietzsche’s ethical ideal, even those that involve overcoming quantitatively significant resistances. For example it should be clear that running a marathon or climbing a mountain would not count. Such activities undoubtedly involve evaluative frameworks in which certain normative standards govern self-assessments, and typically involve high degrees of

³⁹ Pippin makes a similar point about the importance of self-dissatisfaction to self-overcoming. Although his more positive claim that self-overcoming is a ‘self-negating and yet self-identifying and self-affirming state’ that involves ‘tension of the spirit’ seems, to me at least, quite vague (see Pippin 2009: 82).

⁴⁰ Note that ethics is used here, and in the preceding discussion, to suggest something broader than ‘morality’ in the specific sense Nietzsche is opposed to. For discussion see 2.3 below.

sacrifice and hardship. But we would be hard pressed to argue that activities like this involve an overcoming of the fundamental evaluations that constitute our practical identity. This is of course not to deny that all kinds of activities, and achievements within them, are said to have involved self-overcoming without such re-evaluation, for example running the marathon faster. My claim is rather that at least part of what Nietzsche means when he uses this term is something different, and in this sense the account of *Nietzschean* self-overcoming suggested so far is not the same as a common sense conception of self-overcoming, whereby an individual has to overcome many obstacles and is determined and willing to make certain sacrifices in order to achieve certain ends.

Furthermore, we can see how this interpretation, if correct, implies that the meaning of Nietzschean self-overcoming might not be sufficiently captured by the perpetual pursuit of overcoming resistance, that is in terms of those will to power psychology readings I considered in section 2.1.⁴¹ If self-overcoming involves, at least in part, re-evaluative, ethical activity then it requires something more of those engaged in it than can be cashed out solely in terms of overcoming resistances. This more specific emphasis is captured by Nietzsche having Zarathustra say that ‘change of values – that is the change of creators. Whoever must be a creator always annihilates’ (Z ‘On Thousand and one goals’), later reminding us that ‘not around the inventors of new noise does the world revolve, but around the inventors of new values’ (Z ‘On Great Events’).

⁴¹ Nietzsche in fact tells us how the pursuit of endless “becoming”, that is, overcoming resistance *tout court* is problematic; ‘The desire for *destruction*, for change and for becoming can be the expression of an overflowing energy pregnant with the future...but it can also be the hatred of the ill-constituted, deprived and underprivileged one who destroys and *must* destroy because what exists, indeed all existence, all being, outrages and provokes him’ (GS 370).

Moreover, we should not confuse self-overcoming, in the way I have interpreted it, with coming to value something differently, where this merely amounts to becoming aware of some aspect of the object of evaluation of which I was previously not aware. For example, say an individual has a positive evaluation of certain institutions and believes that they are noble in their pursuit of knowledge, yet, after years spent in them comes to realize that the pursuit of material wealth is in fact more important to these institutions. On this basis their valuation might change to a negative one. Note though that the normative standards for making evaluations have not changed. In fact a precondition of this new negative assessment is their holding onto to the evaluative framework in which the pursuit of knowledge for its own sake is praiseworthy and the pursuit of material wealth is objectionable. These are the fundamental evaluations that are not under question in such instances.

More positively, I think the reading of self-overcoming presented thus far also allows us to specify more clearly the ‘self’ that is involved in ‘self-overcoming’.⁴² As has been argued it is our practical identity, as constructed through certain fundamental evaluations, that is revalued in Nietzschean self-overcoming. Such a reading therefore resists any overly literalist interpretation of self-overcoming that might confuse practical identity, as primarily a matter of one’s self-conception, with something like personal identity, which is a matter of psychological continuity for temporally extended subjects. If self-overcoming referred to the latter rather than former, putatively involving a literal death of the self, then the notion would become problematic. Since understanding self-overcoming as something that can be reasonably attributed to me after the re-evaluation, presupposes a psychological continuity that could not be overcome on pain of the self in question after self-overcoming referring to a literally

⁴² The role of the self in self-overcoming is something that I suggested the will to power *psychology* readings do not say enough about, a point evidenced by the kinds of questions it was suggested their accounts generate (see Section 2.1).

different person. Velleman stresses the need to keep this distinction in mind, telling us that ‘when someone suffers an identity crisis, as we call it, what is threatened is not his identity as a person but his conception of himself as a person, which might also be called his sense of identity or his sense of who he is’.⁴³ So, with Nietzschean self-overcoming, interpreted in my way, it is our fundamental evaluations, those that are constitutive of our frameworks for self-evaluation and are essential to our practical identities as self-conceptions, which are re-evaluated.

A good example is the Christian-turned-atheist who achieves self-overcoming in this way by abandoning her faith. Not only would such an individual reject the religious-moral norms that allow for ethical assessments such as ‘good Christian’ or ‘sinner’, but by abandoning her faith the self-evaluative framework in which she made these kinds of assessments is rendered obsolete. Therefore, the overcoming of this self-evaluative framework renders her practical identity as she previously conceived of it destabilized. She can no longer rely on that familiar set of fundamental evaluations and norms to assess her, and others, worth and the former ethical significance of her actions, thoughts, desires and projects, her self-conception, is re-valued.⁴⁴ In praising such a difficult shift, Nietzsche writes in a note that, ‘Heroic = that is the striving for one’s

⁴³ Velleman 2006: 356. Velleman’s target is an idea of ‘motivational essences’ he finds in the work of Frankfurt (1999) and Parfit (1984), according to which if one was to change or abandon particularly important, essential motives, those that are supposedly constitutive of one’s personality, then something akin to literal death of the self would be the result. Whilst I do not have the space to develop this point more fully my sense is that Nietzsche, like Velleman, would most likely want to resist the idea that we have ‘motivational essences’ (see Introduction fn.4) and also Frankfurt’s contention that we have an obligation to self-preservation in these terms (see Frankfurt 1999: 138, 139). In fact the commitment to self-overcoming might be thought to be opposed to what Nietzsche would perhaps see as a kind of ethical conservatism in Frankfurt’s view.

⁴⁴ See Velleman 2006: 357. Tevenar also makes a similar point (see Tevenar 2013: 275).

total destruction into one's opposite, the recasting of the devil as God; that is *this degree of cruelty*' (KGW VII.I.I.67).⁴⁵

Through this example I think we can see that for something to count as an instance of self-overcoming it cannot merely question a self-evaluative framework, but it must also involve a shift away from it. The question of a shift towards what then becomes important, that second desideratum set out at the beginning of this section of the ends of self-overcoming (or what it is an overcoming towards). Although I will hold off on considering this till Section 2.4 since I will be able to engage with it more adequately at that stage.

2.3. *Beyond Good and Evil*

Taking up the exegetical case more directly, I now want to argue that Nietzsche's idea of achieving a standpoint 'Beyond Good and Evil' is an instance of self-overcoming as described above, that is an overcoming of a self-evaluative framework. This section's role is therefore primarily exegetical, showing how a project at the centre of Nietzsche's ethical thought admits of a natural interpretation in the terms set out above. However, in doing so I will also be able to draw out some further significant aspects of self-overcoming as an ethical ideal.

In *Beyond Good and Evil* Nietzsche describes the project I have in mind:

But today, thanks to a renewed self-contemplation and deepening of humanity, shouldn't we be facing a renewed necessity to effect a reversal and fundamental displacement of

⁴⁵ Velleman makes the sound point that whilst, due to the radical nature of the shift involved in such a revision of one's self-conception, a certain resistance to such self-overcoming would seem justified, that we would nonetheless not expect as much resistance as for 'sake of literal self-preservation' (Velleman 2006: 356).

values? Shouldn't we be standing on the threshold of a period that would be designated, negatively at first, as extra-moral...The overcoming of morality – even the self-overcoming of morality (BGE 32).⁴⁶

In order to understand what Nietzsche means in this passage it is helpful to look to *On the Genealogy of Morality* where he provides a more detailed setting for this specific achievement. There we find out that the overcoming of morality involves the rejection of a self-evaluative framework that was born of the most world-historical reversal of values, what Nietzsche calls the 'slave revolt in morality' (GM I 7). Whilst Nietzsche in this text paints a complex picture of this phenomenon in terms of its psychological motivations and historical conditions, for our purposes we can focus on the evaluative dimension of this original reversal.

One of the least interpretatively contested claims about Nietzsche's account of the development of what he calls 'European morality' (GS 343, 380) is that he thinks it begins with the triumph of slave morality which represented an overturning of the self-evaluative framework of noble morality. He tells us it was a 'revaluation of all former values', the triumph 'over all others ideals, all noble ideals' (GM I 8, cf. GM I 16; Nietzsche identifies noble morality with the dominant self-evaluative framework of Greco-Roman culture). Whereas, we are told, noble morality expressed itself by fundamental evaluations whereby 'good = noble = powerful = beautiful = happy' (GM I 7) - what Nietzsche calls the aristocratic value equation – slave morality overturned this in favour of new normative standards according to which:

"Only those who suffer are good, only the poor, the powerless, the lowly are good; the

⁴⁶ Whilst the interpretation of the overcoming of morality set out in this section captures something central to the notion, there is also a different sense to the idea, namely that Judeo-Christian morality contains within itself the resources for its own overcoming (see GS 357).

suffering, the deprived, the sick, the ugly...whereas you rich, the noble and powerful, you are eternally wicked, cruel, lustful, insatiate, godless, you will also be eternally wretched, cursed and damned"! (GM I 7)

With this re-evaluation in mind we can make sense of Nietzsche's description of what the overcoming of morality amounts to when he says 'it has been sufficiently clear for some time what I want...with that dangerous slogan which is written on the spine of my last book, *Beyond Good and Evil*... at least this does not mean "Beyond Good and Bad"' (GM I 17). Nietzsche is clearly guarding against the misleading idea that in overcoming the particular morality he is opposed to we abandon all self-evaluative frameworks. In this sense it is correct to say that the overcoming of morality, and self-overcoming generally, does not seek to abandon self-evaluative frameworks in favour of evaluative nihilism (i.e. a world without any self-evaluative frameworks).⁴⁷ This could not be Nietzsche's position since that *would* mean going 'Beyond Good and Bad' (GM I 17) and as he has Zarathustra say, 'only through esteeming is their value and without esteeming the nut of existence would be hollow' (Z, 'On a Thousand and One Goals'). Rather, Nietzsche is suggesting that we overcome the specific self-evaluative framework associated with the morality that he takes to have come to dominate Modernity, i.e., 'our entire European Morality' (GS 343) which he sees as a development of slave morality's original reversal of the self-evaluative framework of noble morality.

Nietzsche's justification for making this association is complex and contentious, since it is dependent on the alleged continuation in modern European Morality of the psychology of *ressentiment* and the normative agenda of harming the highest exemplars of humanity in favour of the 'herd', largely, he contends, through Christianity or

⁴⁷ Nietzsche explains what his immoralism amounts to in EH IV 4.

sublimated Christian values.⁴⁸ However, with regard to achieving a standpoint ‘Beyond Good and Evil’ it is clear enough what the critical object of this project would be insofar as Nietzsche describes it as the achievement of a ‘point beyond our good and evil...by which I mean the sum of commanding value judgements that have become part of our flesh and blood’ (GS 380). So, it makes sense to find Nietzsche concluding that European morality is ‘only one type of human morality beside which, before which, and after which many other (and especially *higher*) moralities are or should be possible’ (BGE 202), i.e., new and higher self-evaluative frameworks.

So, we can now see, through Nietzsche’s account of the overcoming of morality, how this project can be given a natural interpretation in terms of the account of self-overcoming set out in Section 2.2. Yet, in providing this example we can make two further stipulations. The first is, as argued above, that self-overcoming involves a movement from one self-evaluative framework to another, *not* the abandoning of all ethical frameworks whatsoever, a point captured by Zarathustra’s claim that ‘whoever must be a creator in good and evil...must first be an annihilator and break values’ (Z ‘On Self-Overcoming’). The second is that self-overcoming should be thought of as historically situated. As Robert Pippin states, ‘the conditions described as necessary for...self-overcoming are clearly here historical (dependent on one’s time) and social (dependent, in some way, on the state of a shared social world)’.⁴⁹

In fact there are two ways of taking Pippin’s point. The first would be that self-overcoming is historically situated to the extent that the critical object of such a project could only ever be an established self-evaluative framework of the culture we are a part of. The second, and suggested by some of Nietzsche’s remarks, especially in *Thus Spoke Zarathustra*, is that certain ‘powerful creators’ (in Nietzsche’s idiom) might also overcome self-evaluative frameworks that are, in some sense, of their own making,

⁴⁸ For development of these themes see Poellner 2011: 120-122 and Leiter 2015: 155-177.

⁴⁹ Pippin 2009: 83.

although this project will have to stand in *some* relation to the self-evaluative frameworks of the culture they are a part of. I shall discuss this later suggestion in terms of the idea of continual self-overcoming in Section 2.4.

However, by considering self-overcoming through the example of the overcoming of morality two problems with my account come into focus. Since we might wonder whether the kind of activity I described in 2.2 was too general, such that it is any self-evaluative framework that we are being encouraged to overcome in committing to Zarathustra's ethical ideal. I will call these the *formal objections* since they both involve specifying ways in which my account of self-overcoming seems problematically non-specific.

The first is that there are certain self-evaluative frameworks that we could make a reasonable, perhaps even Nietzschean, case for preserving. We could, for instance, point to the ethical framework of Nietzsche's nobles, who are praised for possessing a self-evaluative framework built around self-affirmations, their fundamental valuations being primarily self-expressions, exhibiting a certainty about their elevated status and showing reverence for everything of a 'high rank' (BGE 262).⁵⁰ The second is that we can envisage situations in which an individual is a committed Nietzschean (whatever exactly that involves) and then decides that Buddhist or Schopenhauerean resignation in fact the right answer to what she now perceives to life's greatest problem, namely that of suffering. Likewise we might think of an individual who was previously a libertine and then converts to being a born-again Christian. The problem is that we seem to have what could be plausibly described as re-evaluations of those fundamental evaluations which structure the practical identities in question, and yet it seems that both instances would, and do, fall foul of Nietzsche's criticisms. In other words, they seem to involve

⁵⁰ See also BGE 287, and GM I 10-12.

Nietzschean self-overcoming, as I am interpreting it so far, and yet fall short of Nietzsche's ideal.

In fact I think these *formal objections* to my account can be met. However to see how requires a detailed consideration of the second aspect of my account, namely the ends of Nietzschean self-overcoming or what is an 'overcoming towards'. Given this I will provide my preferred responses in Section 2.4 after considering two different solutions to the respective *formal objections* that are to my mind unsatisfactory.

One response to the first *formal objection* would be to argue that the self-evaluative framework we assent to in overcoming 'European morality', in shifting away from *those* norms of self-assessment, is actually noble morality. Such that the critical object of self-overcoming would be 'European morality', in order to bring about a re-assent to noble morality. Yet, it should be stressed that any return to the specific self-evaluative framework of the Greco-Roman world is not a possibility Nietzsche seriously entertains; 'we children of the future...we 'conserve' nothing; neither do we want to return to any past' (GS 377, cf. GS 40). This is not to undermine the significance Nietzsche attaches to his reading of noble morality and its Homeric values (courage, heroism, beauty etc.), since this serves an important contrastive function insofar as he believes its fundamental evaluations were not based on *ressentiment* or normative ends deleterious to the highest forms of human excellence. Rather, it is just to stress that Nietzsche's emphasis on 'future moralities' (BGE 202) and 'undiscovered land the boundaries of which no one has yet surveyed' (GS 382) suggests that these new self-evaluative frameworks are not simply those of the Greco-Roman world.⁵¹

Addressing the second *formal objection*, one might argue that that it is not just that the overcoming of morality is an important instance of self-overcoming, but more

⁵¹ As I will try to make clear in Section 4 of the third Chapter the idea of 'nobility' comes to play an important role in Nietzsche's meta-ethical ideal through the exemplar of the Future Philosopher.

strongly that it is only a particular self-evaluative framework, namely ‘European morality’, that Nietzsche wants to overcome. In other words, overcoming the self-evaluative framework of ‘European morality’ would be the *only* instance of self-overcoming that Nietzsche is interested in. Further, since Nietzsche takes European morality to be defined through its history in the ‘slave revolt’ and its continued dependence of the psychology of *ressentiment*, in fact the critical object of self-overcoming could be *ressentiment*-based moralities, including, for Nietzsche, socialism, anarchism, nationalism, feminism, and utilitarianism (GM I 11).⁵² Self-overcoming could then be framed as involving a shift away from self-evaluative frameworks in which ethical assessment of oneself and others are ‘re-touched, re-interpreted, and reviewed through the poisonous eye of *ressentiment*’ (GM I 11). Clearly Nietzsche thinks this psychology, a kind of motivated self-deception which distorts certain evaluative features of the world (primarily qualities of persons) in order to enact an ‘imaginary revenge’ (making up for a kind of powerlessness), plays a crucial role in constructing and maintaining the fundamental evaluations that govern the self-evaluative framework of ‘European morality’, establishing its normative standards of ethical conduct.

Nevertheless, such a position faces exegetical difficulties as a satisfactory account of Nietzsche’s ethical ideal of self-overcoming. Since what then are we to make of his praise of *constant* self-overcoming in BGE 257 and EH I 8? Moreover, we might wonder what it means when ‘life speaks’ to Zarathustra and says, ‘whatever I create and however much I love it - soon I must oppose it and my love; thus my will wills it’ and later in the same passage when we are told that ‘good and evil that would be everlasting – there is no such thing! They must overcome themselves out of themselves again and again’ (Z, ‘On Self-Overcoming’). Elsewhere in *Thus Spoke Zarathustra*, Nietzsche has

⁵² See also GM I 5, BGE 225.

Zarathustra say that ‘good and evil, and rich and poor, and high and trifling, and all the names of values: they shall weapons and clanging signs that life must overcome itself *again and again*...(Z, ‘On the Tarantulas’ my emphasis). In other passages Nietzsche expresses similar thoughts, saying ‘life, to us, that means *constantly* transforming all that we are into light and flame’ (GS PII 3, my emphasis), and describes the way in which ‘again and again we keep coming up to the light, again and again we experience our golden hour of victory, - and there we stand, the way we were born, unbreakable, tense, ready for new, more difficult and distant things’ (GM I 12). Nietzsche even describes Beethoven, a figure he has almost uniform praise for, as ‘somewhere between a brittle old soul that is constantly coming apart and an overly young, future-oriented soul that is constantly on its way’ (BGE 245).⁵³ Moreover, when Nietzsche highlights the notion of some kind of continual self-overcoming in this way he seems to offer it as central to his ethical ideal, as representing some kind of commitment, achievement, or attitude we can take towards our ethical lives and practical identities. So I think the above response, claiming that self-overcoming only relates to ‘European morality’ will not work. Nevertheless, what this commitment to *continual* self-overcoming amounts to remains unclear, and I will now provide a reading of why this concept is sometimes given this form.

2.4. *Self-overcoming as a continual activity*

This section will give an account of what self-overcoming, as an ethical ideal, is an overcoming towards; the second question from the beginning of this section. I think we can provide an answer to this by considering Nietzsche’s idea that we should be committed to self-overcoming continually, and the relation between this idea and the

⁵³ See also GS 382, TI VIII 10.

evaluative malaise he sees humanity confronted with in a post-moral age, that is a future in which European morality is no longer subscribed to and the spectre of the last human being seems a genuine possibility.

However, *prima facie* the idea of *continual* self-overcoming, at least in the way I am reading self-overcoming, seems worryingly arbitrary and perhaps even psychologically implausible. That is to say, can we really think that Nietzsche's ethical ideal involves not just the overcoming of one particular self-evaluative framework (say that of 'European morality'), but a commitment to doing so again and again? As Jaspers notes, we could see such a demand as a 'constant self-crucifixion terminating in nothingness'.⁵⁴ Contrary to first impressions I think we can give a plausible reading of this idea, and moreover understanding what Nietzsche intends by making *continual* self-overcoming a central part of his ethical ideal will allow me to respond to the *formal objections* proposed at the end of 2.3. Also, it is worth bearing in mind that it is in the voice of the exemplar Zarathustra, as representing Nietzsche's ethical ideal, that self-overcoming is most often presented in this *continual* form, such that if anyone is committed to *continual* self-overcoming then it seems like Zarathustra is.⁵⁵

The suggestion that we should pursue self-overcoming continually could be framed as a commitment to the Enlightenment ideal of never-ending re-evaluating and re-assessing. Although, according to my account of self-overcoming, what would always

⁵⁴ Jaspers 1997: 393-394. This objection might be similar to one which Hegel directed against the Romantics ideal of endless change (specifically as found in Novalis and Schlegel's "Irony"), claiming that they were expressive of a 'bad infinity' (see Hegel 1975: 64-69). Yet the Romantic ideal of endless change, as a kind of endless striving due to any finite end once attained then being experienced as unsatisfactory, whilst similar in certain regards is not the same as Nietzschean self-overcoming, which on my reading requires something more specific than this (i.e., ethical re-evaluation).

⁵⁵ In an interesting note Nietzsche writes 'NB: Zarathustra, constantly in a parodic relation to all previous values, out of plenitude' (WLN 7 [54]). As Pippin also rightly highlights, Zarathustra embodies 'a commitment to constant self-transcendence' (Pippin 2006b: xxxi).

be under question in such an activity are the fundamental evaluations and norms of self-assessment through which we construct our practical identities.⁵⁶ Yet, we might wonder how plausible this is. Insofar as it we take it literally, as if in every moment and with every thought we re-assess our fundamental evaluations then it seems far fetched, bringing to fore concerns about whether a subject with such an attitude could think or act at all. For even the most critical self-understanding cannot entertain the suspension of all such values at once. However, we need not take it this way. Rather, continual self-overcoming could make more sense as a stance of *openness*, similar to how we might characterize an ideal scientist's attitude as always being open to revision, improvement, and in some instances wholesale overhaul of paradigms. We would be mistaken to think that the ideal scientific practitioner is constantly, or at one specific moment, overhauling everything, rather what is important is the continual commitment to the possibility of wholesale shift in terms of on-going attention to evidential or theoretical limitations. So, this particular attitude of ethical openness would never allow us to be done with ourselves, it would therefore serve the function of always holding open, as a possibility, those 'future moralities' (BGE 202, i.e., new self-evaluative frameworks).

Nevertheless, what, more precisely, are these new and future moralities that self-overcoming, continual or otherwise, is directed towards? How, that is, can we make more sense of the ends of self-overcoming? By providing an answer to this question we will be able to see the limitations of the reading of continual self-overcoming just presented and suggest an alternative.

⁵⁶ A commitment to continual self-overcoming might therefore seem close to an aspect of Sartre's ethics: 'authenticity must precisely lay claim to live this very situation: this will be love as *tension*...this lived calling into question of self by self...the shifting ensemble of perpetually calling things into question and or perpetually surpassing them' (Sartre 1992: 477- 478).

So, if self-overcoming involves a change in a self-evaluative framework as directed towards so-called ‘future moralities’ (i.e., new self-evaluative frameworks) then we might reasonably want to know more about the latter’s nature and origins. Such supposed new self-evaluative frameworks, despite Nietzsche’s at times misleading talk about creating values, should not be thought to come into existence *ex nihilo*. Rather, I think Taylor describes something that seems close to the ends of Nietzschean self-overcoming:

With these seekers...we are taken beyond the gamut of traditionally available frameworks. Not only do they embrace these traditions tentatively, but they also often develop their own versions of them, or idiosyncratic combinations of or borrowing from or semi inventions within them. (Taylor 1985: 17)

This perhaps captures, whilst remaining non-specific, what we might expect to find in such future moralities. Yet, what should be remembered is that one of the key reasons Nietzsche seems to be drawn to self-overcoming as an ethical ideal is that the ends to which such ‘semi inventive’ projects are directed cannot be exhaustively articulated in advance, or put another way they are what I will call horizontal. Self-overcoming then might be thought to express, as Poellner puts it, ‘the desire for self-transformation through an orientation towards what Nietzsche calls ‘ideals’, the contents of which cannot be determinately specified by us as we are’.⁵⁷ I want to argue that this idea points towards a more distinctive understanding of continual self-overcoming than the stance of ethical openness captures. In order to see this more clearly we need to first make a distinction between two different ways in which we might think of the ends of self-overcoming as horizontal.

⁵⁷ Poellner 2012: 77.

For all I have said so far we could think that whilst, at the initial point of engagement in such re-evaluative projects, these future moralities remain to be specified, that they cannot be articulated in advance by us as we are at present, that over the course of such self-overcoming we would, in the end, come to a fully articulate, no longer inchoate, self-evaluative framework. Like the ideal scientist who overcomes the theoretical or evidential limitations of a certain hypothesis in favour of a new one that is a better fit, we might think that there are no limitations, in theory, on achieving and conceptually formulating this new self-evaluative standpoint. It is in this sense that Taylor remarks how, when engaging in such revaluations, we appeal to some deep ‘unstructured sense of what is important, which is as yet inchoate’, but that in the end this is brought to articulation.⁵⁸ The commitment to constant ethical openness dictates that we could not be complacent, but read in this way it seems that achieving the ends of self-overcoming is theoretically attainable, even if in practice we might find it very difficult to maintain this commitment continually.

Yet, there is a more distinctive way of taking Nietzsche’s point about the horizontal nature of future moralities. We could say, *pace* Taylor, that it is not merely at the beginning of our ethical re-evaluations that the higher standpoint which we strive for is inchoate, but rather that in trying to bring to definition our ethical ideals, by, in some

⁵⁸ Taylor 1989: 41- 42. Ridley explores a different way of framing similar ideas in relation to Nietzsche’s ethical ideal as a kind of particularism modelled after artistic agency or creativity. Drawing on Nietzsche’s description of the way in which artists follow ‘laws that defy conceptual formulation precisely because of their hardness and determinateness’ (BGE 188), Ridley argues, following Kant’s reflections on artistic genius, that Nietzsche’s ethical agent does not have a formula for practice or action in advance which he then might apply. Rather, in a similar way to how an artist ‘sees what he should do’ as it reveals itself precisely by being involved in the practice of completing a musical phrase or completing a painting, as a subtle kind of seeing-doing-and-responding dynamic in the artistic performance itself, Ridley claims that the ‘good man ‘perceives’ what a situation requires of him, even though there is no statable rule that allows him to do this’ (Ridley: 2005a: 214).

sense, making them ‘objective’ and conceptually articulated, we somehow betray them. Such that when we try to move closer towards definitely achieving and articulating this new self-evaluative framework Nietzsche thinks we in fact somehow falsify or fall short of what was more authentic in our original, adumbrated, commitment to something higher. In this sense we might highlight a kind of anti-conceptualist aspect to Nietzsche’s thinking about ethical ends and ideals. For example consider these two particularly revealing passages:

I caught this insight on the wing and quickly took the nearest shoddy words to fasten it lest it fly away from me. And now it has died of these barren words and hangs and flaps in them - and I hardly know any more, when I look at it, how I could have felt so happy when I caught this bird. (GS 298)

We stop valuing ourselves enough when we communicate. Our true experiences are completely taciturn. They could not be communicated even if they wanted to be. This is because the right words for them do not exist. The things we have words for are also the things we have already left behind. (TI VIII, 26)⁵⁹

This theme is complex and when Nietzsche engages with it his prose often becomes more poetic. Yet given the difficulty of trying to conceptualize a point about the problematic nature of conceptualizing our ethical ideals it is not surprising that he often

⁵⁹ See also BGE 97, 160. Katsafanas argues that for Nietzsche there is a distinction between conceptual and non-conceptually articulated mental states that is co-extensive with the distinction between conscious and unconscious mental states. More specifically he claims that Nietzsche thinks that ‘concepts are not primarily designed to portray their objects accurately; rather, they are designed to facilitate human interaction’ (Katsafanas 2005: 17). I expect that the way in which, for Nietzsche, we ‘falsify’ our ethical ideals by conceptualizing them is a particular example of the way in which he thinks conceptually articulated mental states fall short of some non-conceptually articulated content.

reaches for a more metaphorical style. For instance in the final aphorism of *Beyond Good and Evil*:

What are the only things we can paint...only ever things that are about to wilt and lose their smell. Only ever storms that have exhausted themselves and are moving off, and feelings that are yellowed and late...And I only have colours for your afternoon, my written and painted thoughts...but nobody will guess from this how you looked in your morning, you sudden sparks and wonders of my solitude. (BGE 296)

Drawing on these passages we might say for Nietzsche there is a ‘wilting’ that takes place in the activity of striving to conceptualize, and so realize, our highest ideals. Put another way, Nietzsche seems to think that there is a fundamentally self-defeating character to our projects of ethical self-transformation.⁶⁰ Interpreting the ends of self-overcoming in this distinctive way, as suggesting something more theoretically, rather than just practically, difficult about articulating the ends of such re-evaluative activities, might allow us to see how *continual* self-overcoming might make sense as an ideal we could be committed to, and moreover how it can combat the post-moral malaise and stave off the last human being. To see how we need to remind ourselves of what this malaise is, and then examine how the first reading of *continual* self-overcoming, as the stance of ethical openness, fares on this score.

Nietzsche is concerned, as documented at various points in this chapter, by the idea that in overcoming ‘European morality’ we might sleepwalk into a kind of evaluative malaise, more specifically to a post-moral future in which we give up on striving for

⁶⁰ This realization might seem melancholic, and Nietzsche talks about how our insatiable ‘thirst to possess’ our ethical ideals, our constant striving for what is ‘beautiful, strange questionable, terrible and divine’, has made it such that ‘nothing will sate us anymore’ (GS 38, see also GM III 13).

higher ethical ideals, ignoring his call for higher moralities after European morality. Nietzsche therefore thinks we might all too easily fall into the trap of favouring that ‘pitiful contentment’ (Z P3) represented by the last human being (Z P5).

So, if we return to the example of the ideal scientist as exemplary of the stance of ethical openness, we might note that whilst he must be open to revision given new evidence or theoretical limitations, if at a given time all such considerations speak in favour of a particular hypothesis then he would reasonably have to accept it. In this way a commitment to openness should not amount to arbitrary skepticism. Rather, the stance of openness is openness to revision given that relevant reason-based evidence presents itself. Applying this to the ethical context, then, we might wonder whether we can really make sense of a situation in which, having had achieved some proto-noble standpoint, that is to say if our fundamental values are no longer those of European morality, we would then be committed to overcoming *this* new self-evaluative framework as well (whatever exact content the latter has).⁶¹ Since it might seem that if we arrive at this standpoint and find no good reason, *pace* a kind of arbitrary self-skepticism, to again revise these new fundamental evaluations, then the demand for continual self-overcoming would be unreasonable and psychologically implausible. In this sense a commitment to ethical openness should not dictate revision regardless, since it is the reasons and intuitions about the limitations of a particular standpoint that provide a good deal, if not necessarily all, of the motivation for engaging in any particular self-overcoming in the first place. Ethical openness should only commit us to openness to reasons as they are presented, and it does not itself supply those reasons.

Consequently, if even some proto-noble self-evaluative standpoint must be overcome, as implied by the idea of continual self-overcoming, then we would have to

⁶¹ This is really just a way of a rephrasing the first *formal objection* from Section 2.3 As will become clear in the text I do not think *continual* self-overcoming read as the stance of ‘ethical openness’ can meet this objection.

have a reason why. We cannot appeal to the idea this commitment might sustain our interest in ethical life in a post-moral context, since it should not be the case that our end is to ‘affirm life’, to find a way of striving towards new self-evaluative frameworks, and that continual self-overcoming best serves this aim. Not only might this betray Nietzsche’s remarks on continual self-overcoming which implicitly resist this commitment becoming a means to an end, but it gives rise to two objections. (1) There might be all kinds of different or preferable means for achieving this end (‘affirming life’) in a post-moral context and (2) we would have to have an independent argument, *pace* nihilism, that ‘affirming life’ in a post-moral context is actually realizable. So, the problem with reading continual self-overcoming as the stance of ethical openness is that it does not, by itself, give us a plausible explanation of the motivation to overcome some proto-noble ethical standpoint or explain why we should keep striving towards new ethical ideals in a post-moral context.

The more distinctive reading of continual self-overcoming, as growing out of Nietzsche’s anti-conceptualism and the horizontal nature of our ethical ideals, suggests a more plausible response to these problems (and the first *formal objection* from Section 2.3). The motivation for re-engaging in re-evaluative projects makes more sense if we understand, on some level, that we have fallen short and are therefore dissatisfied with what we have ended up with as the result of a particular self-transformative project. In this way my constant striving to overcome myself would not be arbitrary, but might be explained in terms of the necessarily always limited attempt of trying to get closer to my ethical ideal, underwriting what might be required to possess ‘the power [and also the reasons] to create for ourselves our own new eyes and ever again new eyes that are even more our own’ (GS 143). However, such continual attempts at ethical self-transformation would not represent the myth of Sisyphus, of attempting to articulate, failing, then starting again from the beginning. Rather we might think, as Nietzsche

suggests in GS 143, that attempts at articulation could come closer to expressing our ethical ideals. Consequently, each more genuine attempt at articulation could overcome, or at least confront, some of the limitations of previous articulations, and therefore come closer to a more genuine expression. And this is the case even if exhaustive, final realization of our highest ethical ideal is never achieved, and if in continually striving towards such ‘future moralities’ we are aware of the limitations and sense of inadequacy that pervades all efforts at articulation in specific, conceptualized self-evaluative frameworks. In this sense the ‘future’ in ‘future moralities’ (BGE 202) does not merely designate its temporal status, i.e., being in the future, but rather more significantly points towards the horizontal character of our highest ideals.

Viewed in this way we can see how the more distinctive reading of continual self-overcoming might also combat the post-moral malaise more successfully. Since if what I end up as always falls short of what I was after, then this should provide good reasons for me, if I am committed, to take up the project of self-overcoming again. By seeing the ends of self-overcoming as inexhaustible in this way, there will always be the opportunity for re-engaging in such projects, an understanding Nietzsche describes as the realization that ‘the secret for harvesting from existence the greatest fruitfulness and the greatest enjoyment is - to live dangerously...Send your ships into uncharted seas’, (GS 282, cf. GS 124) those ‘ships’ directed towards ‘distant futures not yet glimpsed in dreams’ (Z, ‘On Old and New Law Tables’). Read in this way continual self-overcoming could provide a more successful way of maintaining ethical interest in ourselves in a post-moral future in which the traditional frameworks of European morality are no longer subscribed to and self-evaluative frameworks take on a more protean and individual character, representing a key aspect of the Zarathustra’s project of staving off the potential slip of humanity into an ethical malaise.⁶² As Zarathustra

⁶² See GM III 14.

says, ‘creating – that is the great redemption from suffering, and life’s becoming light. But in order for the creator to be, suffering is needed and much transformation. Indeed, much bitter dying must be in your life, you creators’ (Z, ‘On the Blessed Isles’).

I think this more distinctive reading of the ends of self-overcoming also provides a potential solution to the second *formal objection* I noted at the end of Section 2.3. There it was questioned how my account of self-overcoming could make sense of those cases of the converted Buddhist and born-again Christian, where we seem to have examples of what could be plausibly described as an overcoming of a self-evaluative framework that structures a practical identity, and yet arguably it seems that both instances would, and do, fall foul of Nietzsche’s criticisms. They certainly seemed to involve Nietzschean self-overcoming as I had presented it up to that point and yet surely we would have to say they fall short of Nietzsche’s (and so presumably also Zarathustra’s) ideal.

One response to this particular objection would be to concede that such cases are indeed instances of Nietzschean self-overcoming as defined but that there is some *other* Nietzschean standard that could be applied, external to his ethical ideal of self-overcoming, which could rule them out. Another response might be to deny that they are in fact instances of Nietzschean self-overcoming because the reading I have given was never attempting to be sufficient characterization of this concept. So perhaps then we could appeal to the will to power psychology readings of self-overcoming to try to deal with such examples. However, I think both of these responses would be to some degree *ad hoc* and undermine the import of the reading of self-overcoming presented in this chapter. In fact I think we can do better by drawing on the discussion of the ends of self-overcoming as presented in this sub-section.

As we have seen, at least one aspect of what is involved in the commitment to self-overcoming as a *continual* activity is a recognition of the way in which some final

realization of our highest ethical ideal can never be achieved; as was noted above the “future” in “future moralities” (BGE 202) does not merely designate its temporal status but rather more significantly points towards the horizontal character of our highest ideals. However, whatever else the Christian or Buddhist self-evaluative frameworks involve, they are both aiming towards some finished or final state of rest. For example in the Christian case we have the *contemplatio Dei*, or the ultimate end of man, which Aquinas describes as ‘the complete good which satisfies his desire altogether’,⁶³ and in the Buddhist case the state of Nirvana, that ‘perfect oneness’ which Nietzsche describes as ‘that finally achieved state of total hypnosis and silence...as an escape from every aim, every wish, every action’ (GM I 17). Moreover, such ends are aimed at by following fully codified and conceptually articulated self-evaluative frameworks in terms of truths and normative standards expressed in canonical texts like the Bible, which prescribe the various Religious practices that have to be followed if one is to achieve such states of ‘eternal beatitude’. So, as Nietzsche says about Christianity, ‘the value, meaning, horizons of values were fixed, unconditional, eternal, one with God’ (WLN 11 [226]).

So even if the converted Buddhist or born-again Christian, by virtue of their ‘conversions’, could be said to have achieved a re-evaluation of a self-evaluative frameworks which structured their practical identities, they could not be said to be engaged in Nietzschean self-overcoming *proper* on my reading because of what it is that their projects of transformation are directed *towards*. In other words, what these instances of self-overcoming are aiming at is a once and for all satisfaction in some final state of rest or peace, aiming for what Nietzsche describes as a ‘deep sleep’ (GM I 17). So, in this sense we can see that whilst ethical re-evaluation might be a necessary condition for something to count as an instance of self-overcoming on my reading, it is

⁶³ Aquinas, *Summa Theologia* Ia, IIae, 2.8

not sufficient. Hopefully this shows how the reading presented in this Section can respond to the second of the *formal objections*.

Whilst more could be said about this dimension to self-overcoming, specifically in terms of its reliance on Nietzsche's anti-conceptualism, my aims in this Section have been more limited. I have given an account of the ends of self-overcoming (what it is an overcoming towards), showing that continual self-overcoming need not collapse into arbitrariness or psychological implausibility but rather might represent a novel dimension to Nietzsche's thinking about his ethical ideal, a feature which also allows us to answer some of the objections to the reading considered here. Therefore the hope is that I have provided a plausible account of what is involved in Zarathustra's commitment to self-overcoming as an ethical ideal.

3. Further objections to the Ethical Ideal of Self-overcoming

Having set out my account of self-overcoming I now want to examine two further, more external, objections to the account. The first concerns whether or not Zarathustra's ethical ideal represents a 'revolt against nature', the objection being that it is somehow expressive of a certain kind of attitude of revenge. In responding to this objection I will also explain how Zarathustra is exemplary of a matured understanding of the connection between self-overcoming and nature. The second objection concerns the relationship between Zarathustra's ethical ideal and a different aspect of Nietzsche's ethical thought, specifically his praise of fatalistic views as expressed in the notion of *amor fati*. I will argue that by providing an 'attitudinal' reading of what 'love of fate' amounts to we can see that there is no irreconcilable conflict here.

3.1 *The revolt against nature*

The charge that an ethics of self-overcoming is expressive of some kind of revenge is put forward by Poellner, who tells us that ‘this strand of thought in Nietzsche...evidently manifests itself as a deliberate ‘struggle’ against his – and in his view, mankind’s – ‘most intimate feeling’...It represents an uncompromising revolt against nature – against man as he (supposedly) is.’⁶⁴ Yet, far from being unaware of this objection Zarathustra is shown by Nietzsche to be alert to it, suggesting the need to distinguish his ethical ideal from ‘teachings of revenge’:

that mankind be redeemed from revenge: that to me is the bridge to the highest hope and a rainbow after long thunderstorms...My friends, I do not want to be mixed in with or mistaken for others. There are those who preached my doctrine of life, and at the same time they are preachers of equality and tarantulas...They speak in favour of life, these poisonous spiders, even though they are sitting in their holes and have turned against life, because they want to do harm...(Z, ‘On Tarantulas’, cf. Z, ‘On Passing By’)

Yet, whilst Zarathustra clearly expresses the desire that his ideal should not be misunderstood as involving an attitude of revenge, there is no rationale given here for countering that charge. Framed in Zarathustra’s own idiom, the idea of ‘turning against life’ and ‘wanting to do harm’ seems *prima facie* at least as applicable to an ethics of self-overcoming which is based on a radicalized form of the ascetic dynamic (Section 1) and involves continually re-evaluating the fundamental evaluations that constitute one’s practical identity (Section 2), as to so-called ‘preachers of equality’ (i.e., socialists, anarchists, liberals).

We can get clearer about how we might respond to this objection by paying attention

⁶⁴ Poellner 1995: 265. See also Lowith 1956.

to a claim that is made elsewhere by Zarathustra about the character of the natural world. In the passage I want to highlight ‘life speaks’ to Zarathustra and tells him that, “I am that *which must always overcome itself*. To be sure, you call it will to beget or drive to a purpose, to something higher, more distant, more manifold: but all this is one, and one secret. I would rather perish than renounce this one thing’ (Z, ‘On Self-Overcoming’). We find a similar idea expressed by Nietzsche’s in a later work, telling us that, ‘all great things bring about their own demise through an act of self-sublimation: that is the law of life, the law of necessary ‘self-overcoming’ in the essence of life’ (GM III 27). The central claim that I want to take from these passages is that nature, as referring to a broad range of organic life, as concerned with ‘life and the nature of all that lives’ (Z, ‘On Self-Overcoming’), is said to, in some sense, essentially involve self-overcoming. Whilst Nietzsche in the notebooks of the late 1880’s develops similar sounding ideas in terms of ‘the will to power in nature’,⁶⁵ I think we can interpret this standpoint, as expressed in these published passages, more simply as an anti-inertial conception of nature. We might put this as follows:

Anti-internal conception of nature: phenomena in the natural world are not satisfactorily accounted for if they are described as merely statically maintaining themselves. Rather they should be thought to essentially involve, *inter alia*, growth, change and development.

In trying to get clear about the entailments of such a view we should firstly note that it should not be thought to involve an overtly teleological conception of nature. Such a conception is often associated with Aristotle, and his claims that (a) the natural world is aiming towards certain specifiable ends which are of a pre-eminently rational character,

⁶⁵ WTP 619, 689, 658,

and (b) that the *telos* of nature has a special relation to human beings (and their projects) as the apex of a rational hierarchy.⁶⁶ The anti-inertial conception of nature need not be committed to either (a) or (b). Rather more simply it can just be framed in terms of the view, as Kaufmann puts it, that ‘nature – and life, which is a natural process – do not maintain any status quo. They are continually in motion, striving towards self-transcendence’, regardless of whether such ‘self-transcendence’ is rational in character or has any special relation to human beings.⁶⁷ So according to such a view the natural world could be plausibly be described as exemplifying, *inter alia*, a certain kind of self-overcoming. However, we should note that whatever *this* kind of self-overcoming amounts to it is not the same as the specifically ethical form of self-overcoming I have argued is Nietzsche’s ethical ideal in this chapter (a distinction perhaps motivating Nietzsche’s use of scare quotes for the term in GM III 27).

With this conception of nature in place I think we can give a reasonable response to the objection that Zarathustra’s ethics of self-overcoming involves a kind of revolt against nature. Broadly speaking the objection was that in being committed to this ethical ideal, one is involved in an activity that is not natural or not in line with the natural order of things. More specifically, it is one’s own nature, or how one finds oneself to be, that is being continually revolted against. We can see this at work quite clearly if we remember that in Section 2.4 it was argued that Zarathustra’s ethical ideal is opposed to any state of rest, insofar as one is always dissatisfied with the results of any particular re-evaluative project. So, the objection runs, we would be constantly seeking to, as Poellner puts it, ‘destroy what is ‘merely nature’ in us’.⁶⁸

However, with an anti-inertial conception of nature in place we can respond to this objection. Since, insofar as someone is committed to such an ideal they could be said to

⁶⁶ Aristotle 1962: 40.

⁶⁷ Kaufmann 1974: 260.

⁶⁸ Poellner 1995: 265

be involved in a project which, albeit more ethically orientated, is in some sense also exemplified in the natural world. As such a commitment to an ethics of self-overcoming could be reasonably described as ‘exemplifying a striving that is essentially natural,’ as Kaufmann puts it.⁶⁹ Of course this response only works if the anti-inertial conception of nature turns out to be true. Yet granted that it is at least plausible this seems like a fair response.

Moreover, a further important aspect of Zarathustra’s exemplary role can be drawn out on the basis of this response. Since, it seems by recognizing this point, that is by having learnt something about the natural order of things from ‘life’ (‘life’ speaking to Zarathustra’) Zarathustra is exemplary of a matured self-understanding, and so is ultimately a more successful exemplar, being shown by Nietzsche to have a response to this objection.

This point can be seen more clearly if we note the following. According to a view in which an ethics of self-overcoming is expressive of ‘revenge against nature’ the individual pursuing this ethical ideal would always feel to some extent alienated from the ‘nature’ that he is revolting against it. This attitude was familiar to the religious ascetics in terms of their negative self-assessments in relation to their ideal, and was therefore strongly connected to feelings of guilt and shame.⁷⁰ Yet, by recognizing an ethics of self-overcoming as tracking something essential to the character of the natural world, Zarathustra is shown to subscribe to a different attitude. More specifically, he no longer makes an assessment of self in terms of a framework that sharply distinguishes an ethics of self-overcoming and some stable nature against which one is supposedly doing violence. Although only suggestive this new understanding potentially opens up the space for a more positive attitude towards the ascetic dimension to self-overcoming. We find this at work in Zarathustra’s idea that there is a certain joy to be found in these

⁶⁹ Kaufmann 1974: 260.

⁷⁰ See GM III 15, 16.

projects, ‘the happiness of spirit is... to be anointed and consecrated by tears to serve as a sacrificial animal’ (Z, ‘On the Famous Wise Men’), and in Nietzsche own voice that ‘the most spiritual people, being the strongest, find their happiness where other people would find their downfall...they take pleasure in self-overcoming: asceticism is their nature, requirement, instinct’ (A 57, cf. BGE 224).

3.2 Self-overcoming and *Amor Fati*

The next objection to Zarathustra’s ethical ideal that I want to consider is the apparent conflict with Nietzsche’s idea of *amor fati* (love of fate) which claims that the ‘formula for human greatness is...not wanting anything to be different, not forwards, not backwards, not for all eternity’ (EH II 10). Similar ideas also seem to be at work in his praise of Goethean serenity as a kind of harmony of the soul (WTP 283),⁷¹ and the claim that what defines nobility is a certainty about oneself and one’s elevated status (BGE 287, GM I 10-12).

The reason for thinking that there is a conflict between ‘love of fate’ and self-overcoming is fairly straightforward, since it is difficult to see how an ideal defined by a commitment to continually re-evaluating one’s self-evaluative frameworks could be reconciled with ‘not wanting anything to be different’ (EH II 10). One response would be to point out the fact that the term *amor fati* only occurs a relatively small number of times in the published corpus (GS 276, EH II 10, EH III CW-4, NCW E 1), and in fact

⁷¹ Whilst detailed consideration of Nietzsche’s relation to Goethe is beyond the remit of this chapter it is worth noting a qualification on Leiter’s slightly misleading claim that all Nietzsche’s references to Goethe are ‘uniformly very positive’ (Leiter 2015: 93 fn.4). In fact in a number of key places Nietzsche criticizes Goethe for exemplifying an affective stance that is too ‘restful’ and ‘contemplative’. Claiming in reference to Zarathustra’s ideal that ‘a Goethe...would not know how to breathe for a second in this incredible passion and height’ (EH III Z 6, cf. BGE 198).

only once before Nietzsche's last works of 1888. Compared to the prominent place that self-overcoming seems to occupy across the corpus one could just dismiss the idea as not particularly central to Nietzsche's ethical thought. However, such a response is ultimately exegetically unsatisfactory. Even if the term *amor fati* is not mentioned with any real frequency fatalistic themes do undoubtedly run through Nietzsche's corpus. For example he writes in a late work that 'an individual is a piece of fate, from the front and from the back' (TI IV 6) and similar ideas seem to be at work in the doctrine of the Eternal Return of the Same in *Thus Spoke Zarathustra*; 'I will return to this same and selfsame life, wanting everything to be as it was' (Z, 'The Convalescent').⁷² Given such considerations I think the conflict needs addressing. However, despite first appearances I want to argue below that these fatalistic themes are not in insurmountable conflict with Zarathustra's ethical ideal of self-overcoming.

To see how I think we initially need to distinguish between *amor fati* as an attitude one might have and fatalism as a metaphysical doctrine which holds that, as Leiter puts it, 'whatever happens had to happen...Classical fatalism involves the notion of some sort of non-deterministic, perhaps even non-causal necessity, and in that sense is a rather cryptic view'.⁷³ This distinction might seem obvious enough but it leads us to a number of important further points. The first is that the attitude of *amor fati*, whatever its precise nature, need not imply a commitment to metaphysical fatalism, and certainly does not, in and of itself, provide an argument for a descriptive claim about the way *events* play out in the world. Beatrice Han Pile makes this point when she notes that 'what 'fate' picks out is existential (rather than metaphysical) necessity, e.g. the

⁷² See also D 130, BGE 231.

⁷³ Leiter 2015: 65-66. Without going into details most modern commentators are in broad agreement that cosmological versions of the ER doctrine, as found in the Nachlass, fail to establish any 'metaphysical conclusions' and do not represent the kernel of what Nietzsche intends with the doctrine. For discussion see Clark 1900: 245-250.

perceived ineluctability of what happens to me from a personal standpoint rather than the interconnectedness of all things'.⁷⁴ What this brings into focus is that the object of *amor fati* is one's own 'fate', that is, it is an attitude towards specific aspects of the way one's life has played out. In this sense we might think such a reflexive attitude presupposes that one is adopting a primarily reflective standpoint in the sense of surveying what has already happened. So, we might think that *amor fati* is primarily an exercise in self-interpretation, examining the narrative strands of one's life up to, and at, a fixed point. This certainly make sense of the fact that the idea has a more important place in Nietzsche's semi-autobiographical work *Ecce Homo* than in any of the works which precede it.⁷⁵

Once the above points are made I think we can rule out that idea that *amor fati* implies any kind of ethical conservatism, the latter standpoint being one which Nietzsche accuses Goethe of when he claims that his affective relation to the world involved a 'slackening of the reigns' (BGE 198). Such that we might think that Zarathustra's ethics of overcoming need not be in conflict with *amor fati* insofar as they are not competing for the same philosophical space. One is an ideal of ethical practice, the other is a reflective attitude towards one's life as it has played out. A playing out which could, and in Nietzsche's case ostensibly did, involve a commitment to, and the pursuit of, ethical self-overcoming, as shown by his reflecting, in *Ecce Homo* that his humanity consisted in 'constant self-overcoming' (EH I, 8), a remark this is made without mention of conflict with *amor fati*. In this sense it might be possible to 'love one's fate', and at the same time, like Nietzsche in *Ecce Homo*, have this reflective

⁷⁴ Han-Pile 2009: 258, fn. 72. See also Ridley 2005b: xvii.

⁷⁵ A similar point might be made with regard to the placement of 'Eternal Return of the Same' in *Thus Spoke Zarathustra*, that is at the end of the original three parts of the text.

attitude involve praising of one's own self-overcoming's.⁷⁶ This view of *amor fati* is expressed by Zarathustra when he tells us that:

One has to learn to love oneself – thus I teach – with a hale and healthy love, so that one can stand oneself and not have to roam around...And truly, this is not a command for today and tomorrow, this learning to love oneself. Instead, of all arts this is the most subtle, cunning, ultimate and most patient. (Z, 'On the Spirit of Gravity' 2)

Specifically the idea that 'loving oneself' is not 'a command for today and tomorrow' seems to gesture in the direction of the above interpretation, such that *amor fati* is not supposed to be a normative recommendation for ethical practice, but rather an attitudinal stance one should take towards one's life when adopting a reflective pose. In this sense the attitude seems to suggest a kind of interpretative 'good will' to all of one's self-overcoming's, which is perhaps a helpful attitude to have if one is attempting to better comprehend, as Nietzsche is in *Ecce Homo*, those kinds of achievement.

Concluding Remarks

By providing an account of the exemplar of Zarathustra as representing Nietzsche's ethical ideal I have been able to highlight the kind of first-order evaluative activities (i.e., normative ethics) that I think Nietzsche values most highly, namely those in which we stand to achieve ethical self-transformation by way of overcoming the fundamental evaluations of our self-evaluative frameworks. Moreover, by providing an account of Nietzsche's idea of continual self-overcoming, I described what self-overcoming is

⁷⁶ It is unsurprising then to find that a work like *Ecce Homo* takes a fairly lax stance towards certain facts of Nietzsche's life in order to make certain particular instances of self-overcoming, say that of Wagner, Schopenhauer and Romanticism, seem even more significant.

plausibly an ‘overcoming towards’ in terms of horizontal ethical ends expressed in the idea of ‘future moralities’.

Having now provided in my first two chapters accounts of the critical ideal through the exemplar of the *free spirit* and the ethical ideal through the exemplar of Zarathustra I want to finally consider the meta-ethical ideal. Although it has been argued that an ethics of self-overcoming is the central positive dimension of Nietzsche’s philosophy of value my thesis still requires a more ‘theoretical’ account of (a) the nature of the agents potentially engaging in such Nietzschean projects, and (b) the conception of value underwriting it. Whilst I will hold off any detailed comparisons between the exemplars and their respective ideals until my conclusion the importance of considering the meta-ethical basis for an ethics of self-overcoming seems intuitively warranted. So, I will now give an account the final exemplar of my thesis, namely the Future Philosopher and the meta-ethical ideal this figure represents.

III

The Meta-Ethical Ideal: Future Philosopher as Exemplar

In this final chapter of my thesis I will provide an account of the exemplar of the ‘Future Philosopher’ (hereafter FP) as representing Nietzsche’s meta-ethical ideal. The FP is characterized by Nietzsche as creating and legislating values, he tells us that they ‘reach for the future with a creative hand and everything that is and was becomes a means, a tool, a hammer for them. Their ‘knowing’ is *creating*, their creating is a legislating, their will to truth is – *will to power*’ (BGE 211).

However, making sense of what Nietzschean value-legislation, and its apparent basis in value-creation, might amount to is difficult.¹ This is not least the case because certain aspects of Nietzsche’s theoretical picture of agents seem to undermine some of the ideas, usually drawn from the Kantian tradition, that have been typically appealed to in attempting to make sense of such an activity. More specifically, Nietzsche seems to be sceptical about the role that reflective deliberation, of at least a certain type, can play in determining what evaluative judgements and commitments a person might end up with. For example, as noted in my Introduction when I discussed self-eliminativist readings, he sometimes suggests that the effective causes of many conscious thoughts, feelings and actions, and often those of evaluative significance, are sub-personal or unconscious drives, the true nature and aims of which are at best opaque, sometimes entirely unknowable, to us (D 115, 119, 129, BGE 32, 117, 158, WTP 289, 372). So, as one commentator puts it, ‘one might think that Nietzsche’s theoretical picture of the self

¹ Given this difficulty Jaspers concludes that for Nietzsche, ‘just what creation is remains necessarily indefinite...[it] is never conceptualized’ (Jaspers 1997: 152-157).

stands in conflict with the *formal conditions* for valuing, more precisely for any valuing of the specific kind that Nietzsche is engaged in promoting'.²

It is one of the aims of this chapter to argue that this conclusion is mistaken, however by considering what we might call the natural view of values we can see why it has an initial plausibility. According to this view, as framed by Katsafanas:

To value X is, in part, to make judgements of the form “X is valuable” to guide one’s practical deliberation about principles such as “X ought to be promoted”, and so forth. Valuing X is simply a matter of being committed to certain reflective judgements and acting in accordance with them. (Katsafanas 2015: 163)

Yet, Nietzsche seems to reject key aspects of this, broadly Kantian, theory of value and its model of practical reason. More specifically he rejects the idea that human beings possess a purely rational will which is able to suspend all affectivity, for Kant, ‘inclinations’, and make evaluative judgements on solely rational grounds. Yet it is precisely this capacity which is said by Kant to (a) underwrite autonomous self-legislation, insofar as ‘he [the rational subject] is subject only to *laws which are made by himself* and yet are universal...he is bound only to act in conformity with a will which is his own but has as nature’s purpose for it the function of making universal law’,³ and, relatedly, (b) allows for action in accordance with what is claimed to be reason’s own law (i.e., the Categorical Imperative as ‘*the supreme principle of morality*’).⁴

² Gardner 2009: 8. By contrast if we look to the Kantian tradition we find a more obvious unity of its practical notions of autonomy or self-legislation, and its theory of rational agency. For a modern example of this see Korsgaard 2009, specifically Ch. 4 and 5.

³ Kant 1991: 109-111.

⁴ Kant 1991: 61.

In contrast to the Kantian subject in possession of a ‘faculty of determining oneself from oneself, independently of necessitation by sensible impulses’,⁵ Nietzsche tells us that moralities, that is our systems of value, are just a ‘*sign language of the affects*’ (BGE 187), and that ‘the will to overcome an affect is, in the end, itself only the will of another, or several other, affects’ (BGE 117, cf. BGE 5). Nietzsche’s thinking about the role of the affects in matters of value suggests, as I will argue in more detail later, a route to a positive meta-ethics. However, it also serves a critical function in that any plausible account of what is involved in the FP’s legislative-activity must accord with what Nietzsche thinks human beings are capable of in this domain.

With this setting in mind my aim in this chapter will *not* be to try and understand the FP in terms of a description of some ‘new values’ that this exemplar could be taken to exemplify, that is in terms of some specific list of evaluative judgements or commitments. For the most part Nietzsche does not provide any substantial account of the values that FP’s create or legislate. Rather, in line with the above considerations, I want to explore the *way* in which Nietzsche suggests FP’s value, and by doing so give an account of what I take to be his meta-ethical ideal.

This chapter will be structured as follows. The first section will argue against attributing to Nietzsche a radical choice account of value-creation. Building on this I will provide some reasons for thinking that we can appeal to the content of conscious experience as a source for understanding the nature of value in Nietzsche’s thought (Section 2). I will then examine the role of the affects in Nietzsche’s meta-ethics, aiming to show how the feature of constraint is present in the FP’s legislative-evaluations by providing a detailed reconstruction and defence of the affective-perception model of evaluation (Section 3). Staying with these themes I will then ask whether certain affective-evaluations might be veridical and how this connects with the

⁵ Kant 1998a: A 534/B 562.

FP's 'noble' mode of evaluation (Section 4). By the end of this chapter I will have given an account of the FP as representing Nietzsche's meta-ethical ideal.

1. Value-Creation as Radical Choice

1.1 Responsibility for Values

In trying to understand what legislating values could amount to the most obvious route might be a volitionalist theory grounded in some notion of radical choice. On such a view value-legislation would involve a creative-evaluative act that would bring into existence new values that I then hold myself to. The appeal of this view might be connected to the way it seems to underwrite responsibility for one's own values, since, such an account might say, it is only by circumventing the facticity of history, society, and the pulls and pushes of desires and emotions, that I can choose my values anew, and therefore genuinely be said to be responsible for them. Persons could, on such a view, determine their essence afresh through creating such authentically subjective values, and so overcome what some philosophers have taken to be a kind of evaluative alienation that comes with being constituted by values not of our own making.⁶

We might see aspects of this view in Nietzsche's idea that we should strive to become human beings who are 'new, unique, incomparable, who give themselves laws, who create themselves' (GS 335, cf. GS 290, TI VIII 19). So, perhaps we could understand what Nietzsche calls the FP's 'capacity for responsibility' (BGE 210) though the extent of such essence determining value-creation as based on some notion of radical choice, measuring such an achievement in terms of 'how *far* someone could

⁶ Sartre gestures towards something like this view of evaluative alienation (see Sartre 2003: 291).

stretch his responsibility' (BGE 212) in this sense. Moreover, such radical choice of values as a means to overcoming evaluative alienation might point towards some kind of Nietzschean affirmation of self, such that one could be 'answerable to oneself... [having] *the prerogative to say 'yes' to oneself*' (GM II 3). Such a reading of Nietzsche's ideas on value-creation is not without its adherents,⁷ and it has a certain appeal, especially given the fact the some of Nietzsche's own statements seem to gesture in that direction. In what that follows I will argue against this reading. Firstly, such a radical choice theory of evaluation (hereafter RCT) is incoherent, and fails to underwrite the kind of responsibility for self it aims to secure (1.2). Secondly, we have good textual evidence for not attributing this view to Nietzsche (1.3).

1.2 Critique of RCT

One way of critiquing RCT would be to show that it is committed to an account of agency that is metaphysically problematic, the target most likely being what seems like its assumption of a contra-causal free will. In Chapter 1 I explained Nietzsche's criticisms of such notions of free will and the related conception of moral responsibility (i.e., deserts responsibility). However here I will pursue a different strategy by drawing on Taylor's criticisms of RCT.⁸ Doing so will also allow me to clarify a number of

⁷ See Deleuze 1986: 185 and Macintyre 1981: 26, 129. In terms of Nietzsche scholars who have been tempted to attribute something like this view to him perhaps Nehamas is the closest. I take this to be implicit in his statement that 'the self, even if it is to be at some point discovered must first be created' and that 'Nietzsche seems intent on undermining precisely the idea that there are antecedently existing possibilities grounded in the nature of things or of people' (Nehamas 1984: 174-175).

⁸ See Taylor 1982: 118-121. Taylor associates RCT with something Nietzschean, although wrongly I will suggest in 1.3.

points concerning issues of evaluation and practical reason that might otherwise be overlooked.

To frame the issues we can turn to the *locus classicus* of RCT as found in Sartre's account of a student who comes to ask for guidance in deciding between staying at home to care for his ailing mother or joining the resistance to fight against the Nazi's.⁹ The finer details of the example are unimportant, since all that we need to imagine is a situation in which we draw a blank between incommensurable ethical alternatives, the demands of each being non-contingently in conflict, such that in choosing one way I necessary exclude the other option. What we have is an ethical dilemma and Sartre's suggestion seems to be that in the absence of any criteria by which we might decide, in which he contends both appeal to moral theory and feeling fail,¹⁰ we must make a radical choice:

In coming to me, he knew what advice I should give him... You are free, therefore choose – that is to say, invent... by what authority, in the name of what golden rule of morality, do you think he could have decided, in perfect peace of mind, either to abandon his mother or to remain with her? There are no means of judging. The content is always concrete and therefore unpredictable; it has always to be invented. (Sartre 1948: 38, 52-53)¹¹

⁹ Sartre 1948: 52-53.

¹⁰ Presumably by 'appeal to feeling' Sartre means any prior motivation we might have to choose a certain option that would undermine the terms of the conflict we are faced with.

¹¹ The discussion that follows should not be taken to imply that RCT is the exegetically correct position to attribute to Sartre's views on value and freedom. Parts of these quotes certainly gesture in this direction although there is a subtlety to what Sartre is saying in appealing to the complexity (or 'unpredictable' nature) of ethical life which might be seen as a criticism of the aspirations of certain strands of moral theory, most typically Kantian ones, which tell persons exactly what to do, and claim that once the correct course of action has been determined one can enact it in 'perfect peace of mind'.

Presumably this kind of response is supposed to have a broader significance for RCT. Namely that if we want to secure responsibility for our ethical lives then we must choose our values in the way that Sartre seems to be provoking in the conflicted student, we ‘invent’. However, I want to show that this example in fact ends up showing RCT to be implausible by revealing as incoherent the idea of a choice made with no reference to reasons. This criticism can be broken down into two stages.

The first is that the Sartrean dilemma relies on the positing of evaluative demands the origins of which are not, and cannot be, based on radical choice. Since we can only make sense of dilemmas of this kind if we understand there as being what are, by the subject’s lights, reasons for A-ing or B-ing which are not chosen, or chosen to be entertained, but which external to radical choice present the subject with conflicting demands.¹² In this sense I can neither wish away such demands, nor choose to entertain them when I please, but rather due to their importance for me they constitute the horns of the dilemma I am faced with. However, even if this point is conceded RCT could respond by claiming that it is when we face a conflict between incommensurables that a radical choice emerges as necessary, even if there would have to exist evaluative demands *not* based on radical choice. This response directs our attention back to nature of the radical choice itself and allows for the second stage of criticism.

Returning to the example, say the student radically chooses, or ‘invents’, and decides he must fight for the resistance. We might wonder what kind of decision this really amounts to, since it cannot be based on reasons. He cannot have weighed the options again and decided that, in fact, fighting for the resistance is the right choice, since to do so, as Taylor notes, would ‘not be an account of radical choice, but rather of our coming to see that our obligation lay here’.¹³ So this option is not available. The only alternative seems to be that the student *de facto* finds himself doing A rather than

¹² See Taylor 1982: 119.

¹³ Ibid: 120.

B, such that he is now fighting for the resistance and not at home with his ailing mother. We might attempt to make sense of this as choosing without reasons, however, as Taylor notes, merely finding myself doing A rather than B seems to undermine the distinction between choice and not really choosing at all, such that we might ask ‘do I really choose if I just start doing one of the alternatives’?¹⁴

The central point of the above is that the radical chooser becomes rationally unintelligible as an agent, in that the distinction between things that happen because of him, rather than happen to him, breaks down if all he can offer us is a description of a *de facto* state of affairs, e.g., ‘I am fighting for the resistance’. Since what is typically thought to be required to understand an action as expressive of my agency is, *inter alia*, the possibility of an appeal to reasons that could be expressed in evaluative judgements, e.g., ‘I owed it to the resistance to go and fight’. As Velleman argues, ‘what animates practical thought is a concern for acting in accordance with reasons...if someone wants to do what makes sense, then...he wants to act in accordance with reasons’.¹⁵ So, we might think that if, at the most basic level of radical choice, the student follows no reasons and cannot offer us (without abandoning RCT) these kind of evaluative judgements at all, then by dint of this his ‘choices’ become rationally unintelligible as *his* actions, such that, as Velleman correctly highlights, ‘the sense in which an agent cannot disown his desire to act in accordance with reasons is that he cannot disown it while remaining an agent’.¹⁶

As such the conclusion drawn by Taylor seems like the correct one insofar as ‘the paradox of the theory of radical choice is that it seems to make the universal features of moral experience what we identify as the failing of rationalization, dressing up as a

¹⁴ Ibid: 120.

¹⁵ Velleman 2000: 140-141.

¹⁶ Ibid: 142

moral choice what is really a *de facto* preference'.¹⁷ However, we might also note that Taylor's criticism is not just directed at RCT but also takes as its target Humean accounts of reasons as issuing from internal sentiment-based preferences, that is desires understood as 'pro' or 'con' attitudes and dispositions, which, along with the relevant beliefs, are claimed to rationalize intentional action (e.g., given that I desire x, and believe y-ing is a means to x, I have a reason to y).¹⁸ The thought underlying Taylor's criticism of Humean accounts is that *de facto* preferences, just finding oneself preferring or desiring (and so doing) A rather than B, cannot provide the right (i.e., in some sense 'objective' or 'good') kind of reasons to underwrite the self-intelligibility of our actions as expressive of our agency, that is as supporting ascriptions of what he calls 'responsibility for self'. Such that for Taylor, the Humean cannot properly account for the contrast between the choices, and the reasons given for them, made by a relatively autonomous subject and the behaviour of compulsive subjects who can still be described as having certain desires and beliefs, and so, potentially at least, acting 'according to reasons' given to them on the basis of their *de facto* motivational repertoire. However these are complex issues that I return to in Section 3.

For now, note that if at this final stage RCT was to claim that the choice issued neither from reasons, desires, preferences or anything of this sort then we border on inconceivability. We should simply ask with regard to what then is the choice supposedly made, and in absence of RCT giving us a plausible answer to this question the whole idea of radical choice being not merely a choice, but the paradigmatic example of the most authentic kind of choice, breaks down. In conclusion, RCT, as an account of how values originate, and how we might hold ourselves responsible for them

¹⁷ Taylor 1982: 120.

¹⁸ See Davidson 1980: 13 and for the critique Taylor 1982: 121. Taylor seems to think that RCT might well collapse into a Humean account in trying to make radical choice anything other than entirely mysterious.

or be answerable to them, is incoherent. Despite first appearances the kind of responsibility for our evaluations that Nietzsche wants to cultivate with his idea of legislating values is undermined by an account that renders the subject of these values rationally unintelligible.

1.3 Nietzsche on constraint

Fortunately, given the above, we have good textual evidence for not taking RCT to be Nietzsche's meta-ethical position. This point can be seen partly through the language of constraint that is present in many passages where Nietzsche's discusses evaluative activities:

Artists might have a better sense of smell even in this matter: they are the ones who know only too well that their feeling of freedom, finesse and authority, of creation, formation, and control reaches its apex when they have stopped doing anything 'voluntarily' and instead do everything necessarily... (BGE 213)

Usually, however, there came to these men of fate that redeeming hour, that autumnal hour of ripeness, in which they *had* to do what they did not even "want" to do – and the deed of which they had hitherto been most afraid fell easily and unsought from the tree, as an involuntary deed, as a gift. (WTP 972)¹⁹

The language of necessity and involuntary deeds should persuade us that whatever exactly legislating values amounts to, such values are not thought, by Nietzsche at least, to come into existence *ex nihilo* on the basis of voluntary acts of will.

¹⁹ For other passages which express similar connections between creation and constraint see BGE 118, BGE 203, WTP 928, WTP 954.

Before continuing it is pertinent to note that in the two quoted passages Nietzsche is providing some kind of phenomenology of creative-evaluative experience, that is a description of the qualitative character, or ‘what it is likeness’, of this activity as it presents itself, in first-personal conscious awareness, to a subject. So, Nietzsche is not saying, here at least, that in fact such activity is effectively caused by some deeper psychological mechanisms, for example unconscious drives. Rather, Nietzsche is making a different point in claiming that the experience of ‘creation’ is accurately described as involving feelings of constraint and compulsion, as given analogously in first-personal accounts of artistic creation (BGE 213, cf. HH 155) and in speculative historical accounts (WTP 972, cf. GM I 2). So, given this Nietzsche would have grounds for rejecting RCT on the basis that it misdescribes the character of creative-evaluative experience. However, the relevance of this point, that of understanding value-creation through an account of the phenomenology of creative-evaluative experience, is that it suggests where we might look for the right account of Nietzsche’s meta-ethics. Namely in a more detailed understanding of creative-evaluative experience, such that an accurate description of that experience and an explanation of what it involves will be central to understanding what kind of meta-ethical ideal Nietzsche’s FP exemplifies.

2 The Appeal to the Content of Conscious-experience

2.1 Gardner’s Puzzle

I said above that to gain a better understanding of what value-legislation amounts for Nietzsche we might try to give a more detailed account of creative-evaluative experience. In Section 3 I will argue that drawing on the role of affective states in his

meta-ethics can provide us with such an account. However, I first want to address a worry about appealing to the content of conscious experience in the case of understanding Nietzsche's theory of value.

It was suggested at the beginning of this chapter there seems to be a lack of fit between Nietzsche's theoretical picture of agents, specifically his scepticism about the role of reflective deliberation in matters of values, and much of what is said about the FP's evaluative activity such as their 'conscious implantation of a unity of method' (BGE 210), 'capacity for long-term resolutions' (BGE 212), and the legislative mode of evaluation more broadly. Nevertheless, whatever these notions amount to it seems Nietzsche does not take them to be either (1) practical fictions; the FP's are *not* presented by Nietzsche engaged in some kind of make-believe valuing, rather their 'legislation' is often characterised, as I will explain in more detail in Section 3, in terms of certain kind of responsiveness to objects they take to be valuable (this is often, although by no means always, framed in terms of how they take themselves to be; 'some fundamental certainty that a noble soul has about itself...the noble soul has reverence for itself' (BGE 287); or (2) epiphenomena; the legislative mode of the FP is not framed by Nietzsche as to be explained away by, or reduced to, sub-personal or unconscious forces. Rather, the notions in question seem premised on the idea that conscious states entertained by the FP's are important in some way, and are therefore explanatorily indispensable to understanding their specific mode of evaluation.

Sebastian Gardner makes a similar point:

The subject who values must understand himself – his self – as the *ground* of the values that he affirms...it is part of what it is to entertain and affirm values in the proper non-alienated, explicitly legislative mode, that one's own contribution, the subject's act of *sponsoring*, be understood as constitutive of the 'object' that comprises one's values. (Gardner 2009: 8)

Gardner is claiming that what is essential to the legislative mode is a kind of understanding on the part of the individual, which takes the form of a conscious reflexivity; a kind of ‘sponsoring’ that involves an awareness of my self as the basis of the values I adhere to. However, Gardner is not merely claiming that this reflexive ‘sponsoring’ is explanatorily indispensable to understanding this legislative mode. He also says that the ‘subject’s act of *sponsoring*’, the FP who says “‘That is how it *should* be!’” (BGE 211), must take that very act of sponsoring to *constitute* the value of the ‘object’. This is slightly obscure at present (I will return to issues of the constitution of value in Section 3), but what Gardner is highlighting seems intuitively connected to Nietzsche’s statement that ‘creating values is the true *right of masters* – the originally rare and noble urge to ascribe to yourself a value that comes *from* yourself’ (BGE 261, cf. BGE 43, GM I 2). For the moment we can grant that legislative evaluations seem to involve self-affirmations of some sort, and that it seems that they cannot be made sense of for the valuing-subject independently of the fact that it is the subject himself who is, in Gardner’s terms, consciously ‘sponsoring’ such values.

The motivation for Gardner’s view becomes clearer when he tells us that ‘Nietzschean man must set value on *himself*, not on some psychological structure’.²⁰ This kind of value-constituting self-affirmation is, Gardner thinks, incompatible, on pain of evaluative alienation, with entertaining a third-personal standpoint with regard to the origins of one’s values. Yet, it is not that the entertaining a third-personal perspective on the origins of one’s values, for example as formed from a power-composite of sub-personal drives, is alienating for this legislating-subject *whatsoever*. Rather, what would be alienating according to Gardner is if ‘*in the very act of legislating* and endorsing those values’, this subject took himself as in fact being moved by the ‘causal effects of pre-normative psychological forces’, for example, if the FP

²⁰ Gardner 2009: 9.

believed, in the very act of self-affirmation, this experience to be no more than the upshot of a certain drive taking power in his psychic economy.²¹

Still we might wonder what to make of the conflict between Nietzsche's theoretical picture of agents and the supposedly reflexive, conscious attitudes that characterize the FP's legislative mode, since Gardner's reflections serve only to make this tension more apparent.²² Below I want to argue that the idea of an insurmountable conflict here is unwarranted, and doing so will also allow me to move forward to consider what I consider to be Nietzsche's meta-ethical stance (and then his ideal) in Section 3.

2.2 The role of conscious experience: reflective vs. pre-reflective

We should note an important point about the role of conscious states in Nietzsche's theoretical picture of agents with regard to questions of value. That Nietzsche is sceptical about the role of reflective deliberations in accounting for the origins of

²¹ Ibid: 16. Katsafanas seems to miss this point when he says, 'although I can reflectively judge that I am only experiencing a valuation as warranted because of the peculiarities of my psychology, this judgment need not lead me to stop experiencing the valuation as warranted' (Katsafanas 2015: 191). Gardner's challenge, at least as I read it, is not concerned with what we might take to be case in a reflective judgement, but the nature of what he calls the 'explicitly legislative mode' (Gardner 2009: 16). Katsafanas remarks that 'there is *no* general problem with taking a valuation to be drive induced while simultaneously regarding it as justified. The problem arises only for specific valuations: it arises only when we judge that the contingency of a valuation is in itself objectionable' (Katsafanas 2015: 29). Yet, surely this concedes Gardner's point, in so far as in this 'explicitly legislative mode', in which the conferring of value is supposed to issue from one's own self-affirming reflexive stance, that such contingency would alienate the subject from those evaluations.

²² Gardner's own suggestion that there is fundamental disunity of reason in Nietzsche's philosophy between the theoretical and the practical, such that these opposing perspectives cannot be reconciled, and that in fact Nietzsche is self-consciously entertaining both as a pedagogic lesson about the limits of reason in modernity, seems textually unsupported. And in any case I think we can do better (see Gardner 2009: 21-29).

values, and our continued commitment to them, does not licence the idea that he regards all aspects of conscious evaluative experience as explanatorily or causally redundant. If Nietzsche does sometimes say things that could be taken this way I take him to have overstated his case in order to make what he takes to be legitimate criticisms of the kinds of distortions and deceptions reflective thinking (often in the mould of abstract theorizing) lead to in the domain of value.

This point, and Nietzsche's tendency for overstating his point, can be seen when he tells us that:

The problem of consciousness (or rather, of becoming conscious of something) first confronts us when we begin to realize how much we can do without it...For we could think, feel, will, remember, and also 'act' in every sense of the term, and yet none of all this would have to 'enter our consciousness' (as one says figuratively). All of life would be possible without, as it were, seeing itself in the mirror: and still today, the predominant part of our lives actually unfolds without this mirroring – of course also our thinking, feeling, and willing lives, insulting as it may sound to an older philosopher. To *what end* does consciousness exist at all when it is basically superfluous. (GS 354)

It would be easy to read this passage as an attack of both the causal and explanatory role of consciousness, indeed this seems explicit in the idea that consciousness is superfluous. Yet, if Nietzsche were claiming that *all* conscious states are superfluous on the basis that 'we could think, feel, will, remember, and also 'act' in every sense of term' whilst possessing no conscious awareness whatsoever of these often paradigmatically conscious states, then it would be relatively easy to question his argument. Since we might wonder what it would mean for us to feel something, and yet register no episodic conscious experience of that feeling whatsoever, in fact we might worry that the very concept of an unconscious feeling (never mind unconscious

thinking, or willing) could turn out be a contradiction in terms. To combat such worries we would need a much stronger argument as to why we should attribute qualitative properties, and also some form of intentionality (i.e., what a feeling is a feeling of) to unconscious states. Such an argument is not given by Nietzsche (here at least). Rather what we have are assertions on the basis of a weak inference that moves from the qualitative character of what we do consciously experience to the ascription of similar properties to that which we do not, i.e., unconscious states. So if read as an attack on consciousness *tout court* GS 354 (and perhaps also those passages similar to it, see GS 357, WTP 676, D 119) seems poorly thought out.²³

However, when we attend to the detail on this passage we in fact find that Nietzsche's focus is, with less attention to terminological distinctions than is required, the potentially distorting nature of at least some kinds of reflective, self-objectifying conscious states; those in which we see ourselves 'in the mirror' (GS 354). To repeat an example I used in the Introduction to this thesis (albeit it there with a different emphasis), Nietzsche is often sceptical of the notion of the self as a separate substratum that stands behind its actions, he tells us that 'there is no such substratum; there is no 'being' behind the deed, its effect and what becomes of it; 'the doer' is invented as an after-thought, - the doing is everything' (GM I 13). As the idea of the 'doer' invented as an 'after-thought' might be thought to show, it is through a third-personal reflective state, in which we take our self as an object of inquiry and separate ourselves from 'the doing', that we end up being drawn, Nietzsche thinks, into a Cartesian-style error, where we distort and misconceive the fundamentally world-embodied nature of consciousness. Such that, if we were to describe what is present in experience accurately, at the pre-reflective level, we find no substantial metaphysical self which stands separated, as a 'being' behind the deed' (GM I 13). Regardless of whether one is

²³ For further discussion of problems with the notion of the unconscious where this involves the idea of distinctively unconscious mental states see Sartre 2003: 73-78.

convinced by the particular claims about the nature of the self here the methodological point is an important one that can be reflected back into a more sympathetic reading of GS 354, and ultimately the role of conscious states in Nietzsche's theory of value.

Perhaps the point of GS 354 could be that we can and do 'think, feel, will, remember' etc. without having to abstract from episodic, experiential conscious states to a higher-level self-objectified state. Hence the idea that the 'predominant part of our lives actually unfolds without this mirroring – of course also our thinking, feeling, and willing lives' (GS 354). So the import of the passage on this alternative reading is that if we want to accurately understand our 'lives' then we should, at least 'predominantly', attend to the character of those states as given immediately in first-personal, pre-reflective terms, attempting to make explicit both the specific content and, in more general terms, the necessary conditions (i.e., conditions of possibility) of those states. In doing so there need be no necessary falsification or metaphysically realistic assumptions (say about the existence of Platonic or Cartesian souls or substances). Rather, we might be able to carry out what Nietzsche describes as 'good philology' in the realm of 'inner experience', being able to 'read off a text as a text without interposing an interpretation' (WTP 479). In this note 'interpretation' signifies a *false* interpretation, like Nietzsche contends results from reflective distortion in the example from GM I 13, and one that can be shown to be such by a more transparent and accurately described account of what we might think of as the 'facts of consciousness', therefore holding open the possibility of a *correct* interpretation of conscious experience.

If this is the right reading of the role of pre-reflective conscious experience in Nietzsche's thinking then it might point towards a solution to Gardner's puzzle. He says that 'Nietzsche's theoretical picture of the self...is consistently external, and can be entertained only for as long as the self is viewed from the outside, whence its

discrepancy with the (necessarily internal) practical point of view'.²⁴ So, we might think that we need something that bridges this divide, a mode of explanation that satisfies an explanation (of the origins) of subject-S valuing-V, without dispensing with that, at present opaque, constitutive role that this subjects experience of valuing-V plays. I think an appeal to the role of affects in Nietzsche's theory of value might be able to satisfy this desideratum. Since affects could be understood as providing the suitable object for a naturalistically credible understanding of values ('morality is just a *sign language of the affects*' (BGE 187)) whilst, as will be argued, necessarily referring to the constitutive, or in Nietzsche terms 'creative', role that the first-person perspective plays in valuing. In order to see how we though need to know more about the affects and their role in Nietzsche's meta-ethics.

3 Nietzsche's Theory of Value

3.1 Affects and Values

The central role of affects or affective experience in Nietzsche's thinking about the nature of value is evidenced in passages such as the following:

What is the meaning of the act of evaluation itself...in short where did it originate? Or did not "originate"? – Answer: moral evaluation is an *exegesis*, a way of interpreting ... Who interprets? – Our affects. (WTP 254, cf. BGE 187, 268)

²⁴ Gardner 2009: 11.

Elsewhere he also talks about affects as ‘the mightiest natural power’ (WTP 386) and the ‘greatest sources of strength’ (WTP 931).²⁵ However, exactly what significance this term has for Nietzsche needs a systematic exposition, especially if it is to provide us with an understanding of his meta-ethics and the distinctive legislative mode of the FP.

We should begin with a basic account of what an affect is for Nietzsche, and Janaway provides a good starting point:

...all affects are at bottom inclinations or aversions of some kind. But their range is extensive...he explicitly uses the term for the following: anger, fear, love, hatred, hope, envy, revenge, lust, jealousy, irascibility, exuberance, calmness...Affects are, at the very least, ways in which we *feel*. Many specific instances are what we would call emotions... (Janaway 2009: 52)

So, in at least some significant instances, affects are, for Nietzsche, emotions. As kinds of emotions we might think they typically involve:

- (1) A felt first-personal qualitative character (‘what it is likeness’)
- (2) An intentional object at which they are directed (‘aboutness’)

For example, as the subject of fear I will experience a conscious anxiety, usually along with certain bodily sensations such as increased heart rate and perspiration. In this sense there is something it feels like for a person to experience fear.²⁶ However, such affects are not typically experienced as some experiential state that assails us *ex nihilo*. Rather,

²⁵ See GM III 12 for the role of the affects in Nietzsche’s ideas about knowledge.

²⁶ Notwithstanding the putative existence of ‘unconscious emotions’. For a discussion of ‘unconscious emotions’ and ways of rendering the notion plausible without sacrificing the ‘felt phenomenology’ condition see Deonna and Teroni 2012: 16-18.

most of the time, affects, as kinds of emotions, are experienced as caused by an intentional object that they are about. Again in the case of fear there is characteristically something I am fearful of (e.g., the aggressive individual on the other side of the room). The point is that for the subject experiencing fear, this emotion is always a fear of something. Even if the intentional object is as complex as a hypothetical or imaginary future state, say, fear of a doctor's appointment I have yet to book, or as vague as some inchoate sense of eeriness in an abandoned house, there, is always something the fear is about, its object in this broad sense, which is typically experienced as causing the felt phenomenology.²⁷

That most affects, as kinds of conscious emotions, possess these two features (1) a felt phenomenology, and (2) intentional objects which are, typically, taken to cause that experience, is relatively uncontroversial.²⁸ We might also add (3) emotional experience is typically *pre-reflective* in that most of my emotional life is experienced in a non-reflective, first-personal mode, and that whilst I can, and in certain cases do, reflect on my emotions, reflection itself is not an essential part of most emotional experience. For example, we do not typically need to reflect on, or form judgements about, a certain

²⁷ Note that in certain cases the *actual* cause of the emotion and its intentional object can come apart. For example my jealousy might be about (have as its intentional object) my partner's conversation with another person, yet its *actual* cause might be a certain pathological mistrust I have of my partner. A lot of affective pathologies might have this kind of character. Although typically in such cases this disconnect between intentional object and cause, does not show up, at the experiential level, at least not without suitable therapy.

²⁸ This seems to be commonly accepted in most modern theories of the emotions. See Deonna and Teroni 2012: 1-6. Christine Tappolet talks of emotions being 'world-guided' and says 'they are responses to how things are in our environment' (Tappolet 2012: 210). However this second claim seems ambiguous, since what are we to say about a fear I have when I imagine something, say a future financial state of insolvency. Does this really represent 'how things are in my environment'? I think characterizing emotions as having intentional-objects in the sense described above avoids this potential ambiguity.

conscious anxiety that we experience as caused by an aggressive individual to experience it as an instance of fear.²⁹

Yet, we might ask what exactly the connection between such affective experience and values amounts to. It seems that the thought Nietzsche wants to express in the passage above (WTP 254) is that our values in some way depend on these affects. Ordinary language certainly implies that there is a close connection between emotions and values insofar as many typical emotional responses seem to involve a registering of value. For example, at the most general level, admiration involves apprehending a certain object as of ostensible value, as admirable, likewise disgust involves apprehending a certain object as of ostensible disvalue, as disgusting. In this sense emotional responses seem to involve the registering of evaluative features of their intentional objects, and do so in a way that often appears to underwrite our judgements and beliefs concerning the evaluative features of those objects.³⁰

²⁹ This is not to say that emotions cannot be brought about by reflecting on something. Reflecting on, say, the plight of starving refugees might, in one sense, cause me to be sad. Yet the reflection that leads to this is not an *essential* part of that emotional experience, since surely the plight of starving refugees can be brought to my conscious attention and lead to the emotion of sadness by any number of means other than that of reflection.

³⁰ For the specific ‘cognitivist’ view that emotions should be understood either as, or essentially involving, evaluative judgements see Nussbaum 2001 and Solomon 1976. As will become clear in Sections 3.2 and 3.3 this is not the view I think Nietzsche holds. However since I will not be considering this view it is worth saying something about its two most central problems. We might note that a judgement is a cognitive state of some sort, more specifically it is a thought with propositional content entertained with a certain attitude. Propositional attitudes are typically, due to considerations relating to inferential relations between thoughts, taken to be composed of content that is conceptual. If emotions, as kinds of evaluative judgements, either are, or essentially involve, propositional attitudes then *ipso facto* their content is conceptual, and in order to cognize the relevant evaluative judgement, and so have the relevant emotion, one must be in possession of the relevant evaluative concepts. Yet it seems at least *prima facie* plausible to think that at least some emotional experience is non-conceptual. For example, we might think that babies or animals can experience fear without being in possession of whatever the relevant (typically linguistically expressible) evaluative concepts are (although see

This picture certainly fits with a basic reading of WTP 254 where Nietzsche tells us that the ‘meaning of the act of evaluation itself’ is based in our affects and elsewhere when he says that ‘value words are banners raised where a *new bliss* has been found – a *new feeling*’ (WTP 714), such that when we evaluate, it is our affects that are doing the interpreting. So part of what Nietzsche seems to be suggesting is that our acquaintance with values, that is to say the fundamental way in which we come to experience the world as having evaluative-significance, is through affective experiences. However, there is also a more distinctive claim in the background here, which might be that affects are not merely explanatory psychological tokens for values, which would suggest, perhaps in a Humean spirit, that we could reduce our evaluative-terms to an affective-cum-dispositionalist idiom. Rather, as we shall see in more detail below, they seem to play an important role in not merely acquainting us with, but in fact *constituting*, the domain of value.

However, before examining this more distinctive claim we should ask, are affects, as kinds of emotions, plausible candidates for underwriting suitable conditions of the objectivity of values? A certain strand of thinking might claim that affects represent the antipathy of the kind of constraint that I suggested was central to Nietzsche’s description of creative-evaluative experience (1.3). Since, we might typically think of emotions as expressing our inclinations and aversions, merely subjective preferences or

McDowell 1994 on demonstrative concepts). As a second line of criticism, if emotions are, or essentially involve, evaluative judgements then emotional recalcitrance, that is, for example, experiencing fear when perceiving a spider whilst knowing that it is harmless, seems to involve such subjects in an identical form of irrationality to that of holding conflicting judgements. As Tappolet argues, such a subject would have to be characterized as judging ‘that the object of his or her fear is fearsome, while also judging that it is not’, yet as she notes, ‘whatever irrationality is involved in recalcitrance, it seems to be of a less acute species than what is involved in inconsistent or contradictory judgments’ (Tappolet 2012: 211). The conclusion being that experience of an object as fearsome does not, or at least not necessarily, involve a judgement (see also Döring 2014).

dispositions many of which we might have contingently acquired. As Nietzsche himself states, ‘you still carry around the valuations of things that originate in the passions and love of former centuries’ (GS 57). Presumably it is this understanding of the emotions as passions, a typically pejorative label, that often motivates talk of being a slave to one’s emotions. For example, these seem to be the intuitions which, at least in part, guide the Kantian model of practical reason. As noted at the start of this chapter, for Kant our ability to assess the moral value of an action, and successfully follow reason’s own law (i.e., the Categorical Imperative) is dependent on suspending, at least for the course of rational deliberation, our motivational repertoire as expressed in such affective states, indeed it seems his notion of rational autonomy and self-legislation is dependent on such an ability.³¹

We have good reason to think that Nietzsche rejects this model of practical reason (BGE 187, GM II 6, GM III 12) and we might also think that he is plausibly right in thinking that our affects are the primary *explanans* of values we hold. For example the *ressentiment* subject’s positive evaluation of selflessness is correctly described as motivated by hatred, rather than by his espoused claim to value selflessness because it is good ‘in itself’ (GM I 7-11). Yet, by stressing the role of the affects as the primary *explanans* of our values Nietzsche could be leaving himself open to the charge of advocating a kind of normative subjectivism. Since if values ultimately repose on the arbitrary push and pull of subjective emotions (not to mention unconscious forces such as drives), seemingly devoid of reasons other than ‘subjective’ ones (i.e., I value V because I am drawn towards it and therefore take it as good for me), then this might be thought to undermine the kind of minimal objectivity that many philosophers have thought be required to distinguish evaluation, as involving some kind of constraint, from mere preference. Remember it was on this point that RCT was rejected and

³¹ See Kant 1991: 72, 125.

through which Taylor criticised a Humean understanding of preferences as reasons, insofar as preferences or internal sentiments might be thought *not* give us the relevantly constraining kind of reasons (Section 1; more on this below). This seems to be the skeptical conclusion drawn by Humean-inspired, anti-realist, readings of Nietzsche's meta-ethics, like Brian Leiter's, which hold that Nietzsche's philosophical psychology, based on the explanatory pre-dominance of affects and drives, decisively undercuts 'any objective vindication for his [or indeed any other] evaluative position'.³²

However, *contra* Leiter, this is not how Nietzsche sees at least some affective-evaluative experience. In a note he complains about 'the misunderstanding of passion and reason, as if the latter were an independent entity and not rather a system of relations between various passions and desires; and as if every passion did not possess its quantum of reason' (WTP 388, cf. BGE 284). Elsewhere he criticises 'blind indulgence' of the affects and suggests they be 'habitually sifted by reason' (WTP 928). And perhaps most importantly he contrasts the 'slaves of the mood and desire of the moment' with a capacity for memory, 'mastering of emotions', and recognizing reason (GM II 3).

Reflecting on this final point for a moment, we might note that Nietzsche seems implicitly open to making an important contemporary distinction between two ways in which we can ask about a person's reasons for acting. The question can be concerned, on the one hand, with *motivating reasons*, where what we are interested in is specifying the considerations in the light of which the agent chooses to perform a certain action, and on the other hand, with what Dancy calls *normative reasons* for acting. When we have normative reasons in mind we are interested in whether the person had, in some sense, 'good' reasons for acting. In other words, we are asking whether it was 'worth' doing in the sense that citing normative reasons (on this definition) can give certain actions the

³² Leiter 2002: 146.

quality of being, as Dancy puts it, ‘right or wrong, sensible or unwise’.³³ So whilst we should remember that there are ‘just two questions which we use the single notion of a reason to answer’,³⁴ it seems that Nietzsche could be read as implicitly open to recognizing the distinction between *motivating reasons* for acting (‘desire of the moment’) and a capacity for recognizing reason where this implies the presence of *normative* reasons in Dancy’s sense, something that speaks in favour of the action. It is of course a different matter what the referential status of such reasons are, i.e., what our reasons pick out, as it were. Although as we shall see in more detail in the discussion that follows Nietzsche seems, at least some of time, opposed to a Humean understanding of our reasons for acting as exclusively psychological states of the agent, that is as desires (as Davidsonian ‘pro attitudes’) and beliefs.³⁵ In fact Nietzsche seems to suggest, at least in certain cases, that only the appeal to apparently valuable *objects*, as represented and therefore given through affective experience, can account for and render intelligible our reasons for acting (although these are complex points to which I will return).³⁶

Concluding the main argument of this sub-section then, we might note that affects, at least some of the time, are *not* for Nietzsche merely subjective pulls and pushes, but rather they possess a ‘quantum of reason’, are amenable to some kind of constraining

³³ Dancy 2002: 2. Dancy resists the label ‘justifying reasons’, since he holds, rightly I think, that in certain instances the term might be ambiguous in that I might be able to ‘justify’ my action in one sense, without it in fact having been a ‘good’ or ‘wise’ action, by appealing to why I did it (i.e., motivating reasons).

³⁴ Ibid: 2.

³⁵ Davidson 1980: 13.

³⁶ If this is indeed Nietzsche’s view then it would bear a similarity to that of Dancy who claims that some of our reason giving ‘offers something that seems not to be a state of the agent at all, but a states of affairs...if the reason that motivates is the thing specified on the right hand side of the ‘because’ in ‘I did it because *p*’, many such reasons are apparently not psychological states of the agent’ (Dancy 2002: 15).

influence by conscious thought, and are even described by him as underwriting a more philosophically sound conception of objectivity (GM III 12). So, in order to gain a better understanding of Nietzsche's account of creative-evaluative experience as involving constraint (1.3) we should explain in more detail how notions such as 'reason' and 'objectivity' might come to play a role in our evaluations, and as suggested, we might appeal to the role of the affects as kinds of emotions. Yet unfortunately Nietzsche's own comments above are more declaratory than explanatory, so some rational reconstruction and then defence of the theory I think we can interpret Nietzsche as holding, will be required over the next three sub-sections.

3.2 The Perceptual Theory of the Emotions: An Introduction and the Exegetical Case

Reconstructing Nietzsche's thought in this area Poellner has argued for a distinctive kind of object-intentionality in terms of the relation between our affective experience and values. He says:

We often experience an object, person, or action that we value (or disvalue) as exerting a certain affective pull, an attraction (or repulsion), on us which seems non-contingently connected with the way the object (etc.) itself is. Our affective response in these cases is itself experienced as not merely contingently *caused*, but as *merited* by the object's intrinsic character. (Poellner 2007: 232)

Initially, I want to focus on the first sentence of this quote, holding off on the second that talks of 'merit' (i.e., normativity of some sort) till section 3.3. To begin building up the account, we should note that this view takes some affective experiences to involve direct, non-inferential, representations of the evaluative features of their intentional objects. This mode of evaluative apprehension is supposed to be analogous to the way

in which in ordinary sense perception we directly perceive, for example, colours as intrinsic qualities of objects, i.e., I directly perceive, without intermediaries, this table in front of me, in certain or suitable circumstances, as *objectively* brown, and therefore as available for re-presentation as such on different occasions by myself and others. As McDowell puts it, ‘an experience of something as red can count as a case of being presented with a property that is there anyway – independent of the experience’.³⁷ So, the first claim of this affective-perception model of values is that our emotions, at least in some cases, directly represent and hence involve a perception of, evaluative-features that constitute part of how that object itself is.³⁸ For example, consider my experience of disgust at the sight of a public execution. What I directly perceive, that is what my emotion of disgust non-inferentially acquaints me with, according to this view, is an evaluative characteristic I take to be present in its object, i.e., that the object instantiates the (dis)value of barbarity. In this way perceptual theories claim that emotions represent their intentional objects as possessing certain evaluative-properties in an analogous way to how ordinary sense perception represents its intentional objects as possessing certain colour-properties, they are what Mark Johnston calls ‘disclosures of sensuous values’.³⁹

However, that some emotions involve this kind of object-intentionality, in terms of an affective-perception of evaluative features, should not be taken to restrict our ability to classify similar affective states under typologies, despite their objects being, in one

³⁷ McDowell 1998: 213. As we shall see in the discussion below this notion of ‘objectivity’ needs to be kept distinct from more metaphysically demanding conceptions where what it would mean for something to possess a quality ‘objectively’ would be to possess it such a way that it could be specified entirely independently of any actual or possible experience for a subject, that is independently of its disposition to give rise to certain states.

³⁸ For similar accounts not coming from a Nietzschean perspective see McDowell 1998, Johnston 2001, Döring 2007, Döring 2014, Tappolet 2012.

³⁹ Johnston 2001: 182. Although see Döring 2014 for a different kind of perceptual view which denies that emotions are like sensory perceptions. See also Todd 2014 for a recent critique of this way of thinking about evaluative features as being present in emotional experience.

sense, particulars (rather than say general evaluative propositions like moral rules).⁴⁰ So, the fact that in one instance my fear is about an aggressive dog, and in another about impending financial insolvency, does not stop me identifying both as instances of fear. Following recent discussion in philosophy of the emotions we might individuate emotion-types by reference to their *formal* object, a notion that we could understand, according to the above view, in terms of the representation, and therefore apprehension, of the same evaluative property being present.⁴¹ So for example, a Rembrandt painting, Nietzsche's prose, Homer's Odysseus, all might be the particular objects of my admiration. Yet, at least according to the affective-perception view, what individuates each affective response as of the type 'admiration', what each has in common, is that they all involve a representation of the particular object under the aegis of the relevantly similar evaluative property (e.g., admirable) which is taken to cause the affective experience.

Note that this view of emotions, as perceptions of value, runs counter to the Humean view of emotions as possessing some kind of 'original existence' that 'contains not any representative quality'.⁴² For example in the case of fear, the object would, for Hume, not be said to instantiate any intrinsic value-properties. Rather, such evaluative properties are claimed to be akin to Lockean secondary properties, that is merely 'perceptions in the mind' that we project onto objects, 'gilding or staining all natural objects with the colours borrowed from internal sentiment'.⁴³ It is this line of thought which leads Hume to claim that in order to truly understand the origins (and presumably therefore the meaning) of evaluative-terms one has to turn 'reflexion into your own

⁴⁰ See Pelsner 2014: 109.

⁴¹ See Deonna and Teroni 2012: 40-41.

⁴² Hume 1978: 415.

⁴³ Hume 1987: 294. See also Stroud 1993 for discussion of Hume's view in these terms.

breast',⁴⁴ such that our emotions, and their evaluative content, can be specified at the personal level without reference to representational content as given through their intentional objects. This view of the emotions, and our acquaintance with value, as we see shall in more detail throughout, cannot account for central aspects our affective-evaluative experience.

Yet before considering and defending more of the detail of the affective-perception view I want to ask whether it is plausible to attribute something like this view to Nietzsche. A central claim of this view is that just as we can only come to apprehend colours through our sense experience of certain coloured objects, likewise, we can only come to apprehend values through our affective experience of certain evaluative objects. In this way we might claim that just as a world in which there was no sense experience would be one in which we were fundamentally unacquainted with colours as we understand them, likewise a world in which there was no affective experience would be one in which there was no acquaintance with values.⁴⁵ As Johnston puts it, 'if one has never been moved or *affected* by the determinate ways in which things are beautiful or charming or erotic or banal or sublime or horrific or appealing, then one is ignorant of the relevant determinate value'.⁴⁶ Yet, Johnston's thesis might appear to be primarily epistemological, claiming that without the relevant affect we would lack access to, be 'ignorant of', the relevant value. In fact I think Nietzsche holds a stronger 'ontological' thesis, namely that affective-perceptions *constitute* the domain of value, such that it is not merely that in absence of affect we are ignorant of the relevant value but rather that in absence of affect the value would *also* lack 'being'. I will call this the *constitution of*

⁴⁴ Hume 1978: 468-9.

⁴⁵ See Tappolet 2012: 222, fn.6.

⁴⁶ Johnston 2001: 183.

value thesis: affective-perceptions *constitute* the domain of value (or if you like, affective experience is a condition of possibility for value).⁴⁷

We can see the exegetical plausibility of attributing this view to Nietzsche by noting, in general terms first, that he is partial to the ocular analogy in the case of evaluative experience, for example, he says that ‘it cannot be doubted that *all sense perceptions are permeated with value judgements*...each individual colour is also for us an expression of value’ (WTP 505) and claims that ‘as soon as we see a new picture, we immediately construct it with the help of all the old experiences we have had’ such that ‘there are no experiences other than moral [i.e., evaluative] ones, not even in the realm of sense perception’ (GS 114). Nietzsche also talks about ‘optics’ of value, describing the ‘faulty optic of the theological’ in whom all value-judgements are supposedly inverted (A9), says that Christianity is way of seeing ‘that cannot be approached with reasons and refutations...you do not refute an eye disease’ (CW E) and makes an equation between ‘affects’ and ‘eyes’ (GM III 13). More strongly drawing the connection between affective-experience and values he claims that ‘value feelings’ involve being able to ‘gaze with many eyes’ (BGE 221, cf. BGE 186), and in other notes talks, albeit with cultural-historical concerns in mind, about ‘negative emotional valuations’, lamenting that ‘our feelings, as feelings about values, are not up to date’

⁴⁷ In this sense Nietzsche’s view is different from those perceptual theories of the emotion that are committed to a stricter, seemingly metaphysical, form of value-realism. See specifically Pelser 2014:110. Rather Nietzsche’s ‘value-realism’ as I am reading it is closer in terms of the status it attributes to values to McDowellian secondary qualities as phenomenal properties, the ascription of which to an object cannot be made sense of for a subject apart from ‘the object’s disposition to present [in certain circumstances] a certain sort of perceptual experience’ (McDowell 1998: 212). So we might say the ascription of the value-predicate ‘admirable’ to a painting cannot be sense of for a subject apart from that paintings disposition to present, in certain circumstances, the appearance of a certain kind of aesthetic excellence, an appearance registered in certain affective experiences of admiration. Although as section 3.3 will argue this affective-evaluative experience also involves an apprehension that one’s response is *prima facie appropriate* (merited by) its object.

(WTP 101). Finally, he claims that ‘every ideal presupposes *love* and *hatred*, *admiration* and *contempt*. Either the positive emotion is the *primum mobile* or the negative emotion’ (KGW VII.2.10.9).

Moreover, reading Nietzsche as committed to claim that affective-perceptions constitute the domain of value (the *constitution of value thesis*) gives us a philosophically plausible explanation as to why he often talks about human beings as ‘givers and granters’ of value (GS 301). We might think the often puzzling emphasis that Nietzsche’s places on the ‘creative’ role human beings have in evaluative matters could be a means of highlighting, amongst other things, the *constitutive* role that our affective experiences play in both our acquaintance with, and the continued existence of, evaluative features in the world, that is as the ‘*primum mobile*’ in our values (KGW VII.2.10.9). After all Nietzsche tells us that what we call ‘life’, that is our evaluatively laden *life-world* which involves ‘assessing, preferring, being unfair, being limited, wanting to be different’ (BGE 9), and without which (without valuing) ‘the nut of existence would be hollow’ (Z ‘On a Thousand and One Goals’), has its foundation in the affects (BGE 258). In other words, if human beings were to cease to have affective experiences there would not remain a realm of value specifiable independently:

Whatever has value in the present world has it not in itself, according to its nature - nature is always value-less - but has rather been given, granted value, and we were the givers and granters! Only we have created the world that concerns human beings! (GS 301)

This reading of GS 301 points towards what we might think of as a *phenomenal* value-realist construal of Nietzsche’s meta-ethics in terms of the idea that value is constituted by human affectivity.⁴⁸ Such a view can be contrasted with and opposed to two different

⁴⁸ See Poellner 2007: 227-261.

standpoints. (1) A global error theory (a global anti-realism) about value, which for those who wish to attribute some such view to Nietzsche either makes him thoroughly skeptical about according any privileged status to any evaluative claims (i.e., Leiter's reading),⁴⁹ or for some points towards meta-ethical fictionalism through the invention of make-believe values as a way forward.⁵⁰ (2) A stronger form of *metaphysically* objective realism about value associated with (a) a view of values as Lockean primary qualities of objects (entirely conceptualizable independent of human sensibility),⁵¹ and (b) a kind of Platonism which would locate the putative mind-independence of values not just at the phenomenological level of intentional experience but at the metaphysical level as putatively free-floating non-natural objective value-properties (e.g., the form of 'beauty'). For example in the Symposium we are told that the form of beauty 'is, in the first place eternal; it does not come into being or perish, nor does it grow or waste away...it exists on its own, single in substance and everlasting'.⁵²

⁴⁹ See fn. 51.

⁵⁰ See Hussain 2007: 157-191 and Reginster 2006: 85. What underscores Hussain's mistake in attributing the idea of 'make-believe' valuing to Nietzsche on the basis of a 'global error theory' about value is, to my mind, a failure to distinguish between Nietzsche's opposition to *metaphysical objectivism* about value, that value resides in noumenal or metaphysical realms (e.g., Kant, Plato) and Nietzsche's commitment to the idea that the world that concerns human beings is valuable precisely in virtue of the constitutive role that human beings play as the 'givers and granters of' value (GS 301). The idea that because we cannot attain metaphysically objective value (say in God or transcendent realms), that therefore *nothing is of value whatsoever* (the error theory; all evaluative claims are false), is clearly an unwarranted inference, and one which Nietzsche often associates with a kind of nihilism which sees the devaluation of the 'highest values' as undermining the possibility of *any* value in the world whatsoever (see TI, 'How the "True World" Finally Become a Fable', WTP 8, 12). That Nietzsche wants to overcome such scepticism and argue for the indispensability of a realm of human value, authenticated by the constitutive role human beings play, seems more exegetically plausible and philosophically promising.

⁵¹ For critical discussion see McDowell 1998: 211-212.

⁵² Plato 2008: 211a.

So we might think the exegetical plausibility of attributing something like the affective-perception model of values to Nietzsche is evidenced by the way in which it joins together what have often been thought to be two central, and related, strands of his meta-ethics. (1) Nietzsche's supposed anti-realism about value can be understood here as a denial of the existence of *metaphysically* objective value in that nothing has value 'in itself' or as he puts it the 'Good in itself' is an invention (BGE P), which is, on this reading, primarily a kind of anti-Platonism. This Platonism was by no means trivial to Nietzsche's critical concerns insofar as it was re-asserted and re-imagined, at least according to him, in various ways both by Christianity, since 'Christianity is Platonism for the 'people'' (BGE P), and his foremost influence and critical target Schopenhauer.⁵³ And (2) the 'creative', we might say 'existential', role human affective experience plays in constituting value. In this sense there *is*, for Nietzsche, value in the world, namely the value we 'give and grant' through our affective relation to that world (GS 301). As Zarathustra says, 'esteeming is creating...esteeming itself is the treasure and jewel of all esteemed things. Only through esteeming is there value, and without esteeming the nut of existence would be hollow' (Z 'On a Thousand and One Goals'). So we might conclude it makes good sense to read Nietzsche at times (and by no means consistently in all passages) as gesturing towards something like the affective-perception model of values and the *constitution of value thesis*.

However, exegetical issues aside, can we not question if it is really the case that value depends in this 'existential' sense on affective experience, and specifically on emotions? After all, we might conjecture that 'higher' cognitive evaluative phenomena, like evaluative judgements and beliefs for example, can be, and in certain cases typically are, held in the absence of the relevant affective experience. For example, say someone reminds me of a novel I once read, and I respond by stating, "it was terrible, I

⁵³ Schopenhauer 1969a: 169-177.

hated it". It seems that I can make such evaluative judgements and do not have to rekindle the experience of dislike I had towards it.

Yet, we need to pay careful attention to what exactly the *constitution of value thesis* is claiming. The claim is not that every propositional thought that contains some kind of evaluative content, like evaluative judgements and evaluative beliefs, must necessarily be consciously co-present with the relevant affective state, which indeed seems an implausible position to hold (although presumably *some* are). Rather, the view need only claim that such cognitive evaluative attitudes must have, at some stage, been dependent on a more primary, typically pre-reflective, affective-evaluative response, to make sense of what their (the cognitive attitudes') propositional content refers to. In other words the *constitution of value thesis* must hold, as a possibility, that we always could, perhaps through remembering or imagining, trace those 'higher' cognitive evaluative phenomena back to a more fundamental affective-evaluative engagement with their intentional objects in precisely those cases where the affective dimension is not present. So, whilst I can 'coldly', that is in absence of the affect, state my evaluative judgement about the novel, this judgement only makes sense, that is, is only intelligible to me, by virtue of the fact that when I originally read it I did indeed have the experience of dislike, and that it was my experience of this object in this affective-evaluative mode that provided me with reasons to form judgements, beliefs and more generally respond to it in various ways (e.g. throwing the book away or resolving to not recommend it).

In this context it is interesting to reflect on what relation we might have to evaluative judgements were we somehow, for example by some form of memory repression, deprived of retrieving any sense of our original affective-evaluative response to the object of that evaluative judgement. The intuition guiding the *constitution of value thesis* is that such judgements would simply strike us as unintelligible. As Johnston

writes, ‘when affect collapses, the immediate intelligibility and appeal of what one is doing is lost’.⁵⁴ That is to say in absence of the availability, in some mode, of the relevant affective aspect of our evaluative experience, we might quite quickly become alienated from certain parts of our evaluative life. Julien Deonna and Fabrice Teroni articulate a similar point, telling us that:

Categorizing an object as funny or shameful is indeed hardly detachable from the understanding that its properties give one reasons to favour or reject it. And we might wonder what sort of understanding of there being reasons to favour or reject an object we would preserve, were we deprived of the relevant emotions. (Deonna and Teroni 2012: 122-123)

In this sense our more cognitive evaluative practices (e.g., judgements, beliefs etc.) are plausibly thought to be dependent on our (often pre-reflective) emotional engagement with the world. A subject deprived of this kind of affective-orientation would, we might think, become inarticulate to themselves and to us *qua* evaluator, i.e., not properly engaged in our evaluative practices, as we understand them. To use Zarathustra’s words, the ‘nut of existence’ might well become ‘hollow’ for them (Z ‘On a Thousand and One Goals’).

3.3. Prima Facie Appropriateness and Rational Assessments of the Emotions

Returning to the finer detail of the affective-perception view, if we accept the idea of some affects as kinds of perception which constitute our fundamental acquaintance with values we can now attempt to understand a further distinctive claim of this view.

⁵⁴ Johnston 2001: 193

Namely that there is a kind of constraint to be found in this view's ideas about object-intentionality that underwrites a sense of normativity. That is, a kind of constraint that we saw to be essential to Nietzsche's description of creative-evaluative experience (1.3) and so essential to building up a satisfactory account of his meta-ethical ideal (I touched on these themes above in considering the problem of the rational intelligibility of affectively 'cold' evaluative judgements, beliefs etc. but now want to make the points more explicitly).

To elucidate this claim we can turn our attention back to the second half of Poellner's original statement which says that 'our affective response...is itself experienced as not merely contingently *caused*, but as *merited* by the object's intrinsic character'.⁵⁵ The claim is that the perception of those evaluative features as present in the object are experienced as not just causing my disgust but *also* underwrites that affective response being experienced as somehow appropriate to its object. We can get clearer about this claim by framing it in contrast to the Humean picture.

We might say my 'disgust' at the execution can be experienced as the appropriate response only insofar as the object which causes it is non-contingently taken by me *as itself* instantiating the evaluative feature of 'barbarity' (which is a disvalue for me) rather than merely 'contingently' doing so in virtue of some sentiment I happen to have acquired and am projecting (i.e., Hume's view).⁵⁶ In this way we might claim that at least a good number of my emotional responses would cease to possess what Johnston calls 'ready intelligibility',⁵⁷ that is would cease to be experienced as what I will call *prima facie* appropriate to their objects, which surely a good number of them are,

⁵⁵ Poellner 2007: 232. This is clearly reminiscent of McDowell's claim that a value 'is conceived to be not merely such as to elicit the appropriate attitude, but rather such as to *merit* it' (McDowell 1998: 221).

⁵⁶ See McDowell 1998: 210-226 and Johnston 2001: 183-185 for a more systematic critiques of the projectivist view of value and affect.

⁵⁷ Ibid: 181.

without me taking those represented evaluative features to be intrinsic characteristics of their objects. So, understanding emotions as registerings of value, a point the Humean is committed to,⁵⁸ seems to have to involve reference to those evaluative features *as* features of their intentional objects *if* at least some of our emotional responses are to be experienced as *prima facie* appropriate. As McDowell puts it, ‘we make sense of fear by seeing it as a response to objects that *merit* such a response, or as the intelligibly defective product of a propensity towards responses that would be intelligible in that way’.⁵⁹

So, connecting this point with meta-ethical considerations, we might say that what it is to value V on this Nietzschean model is, in the basic case, to both have the relevant affective response of favouring or disfavouring and at same time experience this response as *prima facie* appropriate to evaluative features I take to be characteristics of the object. For example, I am not merely attracted to an individual, but I experience my attraction as *prima facie* appropriate in light of the fact that I take her to instantiate certain evaluative qualities, say beauty of some sort.

However, Katsafanas has raised an objection to this theory of value, claiming that ‘many agents have values but give no thought to their justificatory status...[so] agents can have values without thinking about whether the values are justified’.⁶⁰ The objection is that this account, and similar ones,⁶¹ are cognitively overburdened in the idea of *justification* constituting part of what it is to value, in a way that is hard to square with Nietzsche’s utterances.⁶² However, I think Katsafanas simply

⁵⁸ See Mackie 1977: 31-5.

⁵⁹ McDowell 1998: 221.

⁶⁰ Katsafanas 2015: 178-179.

⁶¹ See also Clark & Dudrick 2007: 192-226.

⁶² One reply might be to argue that it is only a rarefied subset of agents, maybe Nietzsche’s ‘higher types’ or FP’s, who value in this way. So this account of value would not need to capture affective-evaluative experience generally but just that which we believe, or more aptly

misunderstands what conception of so-called ‘justification’ is necessary to this theory of value. It overlooks the point that the ‘justification’ in question, or more aptly the appropriateness of an emotional response to an evaluative-object, can, and often is, given pre-reflectively as a *prima facie* appropriateness. Insofar as the appropriateness is given in this non-inferential mode, as part of the immediate affective experience, then it does *not* involve reflective *judgements* about whether or not the response is appropriate. In this sense there is what we might call a *primitive* normativity at work in pre-reflective affective-evaluative experience.⁶³ Let me explain this response and the conception of *prima facie appropriateness* as a kind of primitive normativity in emotional experience in more detail.

The affective-perception view of values should *not* be read as claiming that I have some affective ‘favouring’ experience, and then in order for this to count as an instance of valuing I then have to entertain separate justificatory thoughts.⁶⁴ For example, according to this view, I do not typically experience a piece of music as merely exerting an affective pull, after which I then separately reflect on whether my response was appropriate in the sense that *judging* so would putatively complete and confirm my experience as an instance of valuing, e.g., my saying “that affective pull was actually an admiration for the beauty of the music”, whereby such justificatory thoughts acquaint

Nietzsche believes, to be exemplary (see Clark & Dudrick 2007: 192-226). However, this seems to my mind unsatisfactory, after all if the issues are meta-ethical ones about the constitution of value we should at least attempt to capture the character of all (or most) affective-evaluative experience. And as will become clear the ‘exemplary’ dimension to the FP’s can be understood as relying on quite a specific way of applying this affective-perception view in a particular circumstance (see Section 4).

⁶³ See Ginsborg 2011. This claim is opposed to Deonna and Teroni who say that ‘appropriateness is not a part of the emotional experience but requires a judgement about it’ (Deonna and Teroni 2012: 101).

⁶⁴ If this were what it was claiming it would be much closer to cognitive theories of the emotions which claim that emotions either are, or necessarily involve, evaluative judgements (see fn. 33).

me with its beauty and my valuing beauty. Rather, what the affective-perception view claims, at least as I am reading it, is that there is a *prima facie* appropriateness in at least some of my emotional responses that is given as a constitutive part of the original affective-evaluative experience at the pre-reflective level, as a kind of primitive normativity. In this way, the ‘reasons for responding’ present in the original experience are misunderstood if they are thought of as external conditions of justification articulated in a reflective mode, separate from some more basic experience describable without them.

That is not to say that I cannot entertain *separate* justificatory thoughts after the affective-evaluative experience. I can reflect on, for example, whether my experience of ‘anger towards S’ can be *judged* to be appropriate (we shall see how this plays out in assessments of emotional responses below). Yet, regardless of any conclusions drawn separately in reflection, when I was experiencing this affect it was in the light of my perception of those evaluative features, S’s obstinacy say, that my anger towards S was experienced as *prima facie* appropriate to its object, experienced as what Sabine Döring calls an ‘appearance of truth’, which is part of the pre-reflective emotional experience regardless of whether or not I ‘would *affirm* the truth of [my] emotion’s content in judgement’.⁶⁵ Perhaps this could be part of what Nietzsche means by his statement that every passion possesses its ‘quantum of reason’ (WTP 388), and earlier in the corpus that ‘a drive to something or away from something divorced from a feeling one is desiring the beneficial or avoiding the harmful, a drive without *some kind* of knowing evaluation of the worth of its objective, does not exist in man’, where the ‘kind of

⁶⁵ See Döring 2014: 133. In this sense Döring seems right to stress that ‘emotions present evaluative features of the world *in a different mode* that value judgements do’ (ibid: 133). It should therefore be clear that on this reading the affective-perception model of values is clearly distinct from the ‘cognitivist’ theories of the emotions like Nussbaum 2001.

knowing evaluation' would be (a) pre-reflective and (b) non-inferential (HH I 32, my emphasis).

Moreover I think this understanding of a kind of primitive normativity as present in affective-evaluative experience is evidenced by the way in which we typically articulate our emotional responses to evaluative objects in reflective judgements. Say, for example, a painting enraptures me, and someone was to ask me "why do you admire it". My explanation, as an articulation of my emotional response, would seem somehow hollow if I was to say, "it just drew me". Rather, in many cases my response to such a question involves making explicit, if I am sincere, what was pre-reflectively implicit in my experience, namely those *prima facie* appropriate reasons for my admiration (its beauty say, or in Nietzsche's terms the 'worth' of the object), which, according to the view I am defending, I experienced originally in an evaluative way and therefore will be explicated in terms of that perceived evaluative feature. The fact that I do not generally struggle to articulate these evaluatively laden reasons when questioned in this way suggests that they were part of the original experience itself.

Yet this is not to say that by making such pre-reflective, *prima facie* appropriate reasons explicit I do not run the risk of distorting them or deceiving myself about them. Since we might think that honest and self-transparent, reflective articulation of our emotional lives, and the reasons we are acquainted with in those emotional experiences, is often disquieting. For in many cases such emotional honesty and what it reveals about what we value (perhaps rather than what we reflectively claim to value) might generate disturbing results that many people might seek to avoid by self-deception and reflective distortion. For example Nietzsche's *ressentiment* subject has a jealous hatred of the nobles which leads him to engage in a self-deceptive, reflective disavowal of a value, i.e., nobility and power, even though his hatred is, at pre-reflective level, premised on

recognizing those values *as* goods, albeit goods he does not exemplify, but which necessarily underwrite his jealous hatred (see GM I, 10-13).

Furthermore, given the above, we might think the perceptual model of the emotions, by carving out a space between my affective experiences and the intentional objects they are about in terms of those object's evaluative features, allows for assessments of emotions in two important ways. The first is through an assessment which takes as its object that *prima facie* appropriateness I have been discussing, asking, in the form of a judgement, if what I experienced as *prima facie* appropriate reasons for my emotional response are indeed appropriate reasons. It is worth noting that by seeking to assess our emotional responses according to certain rational standards we have to take those responses (and their intentional content) themselves as objects and so leave the more immediate plane of pre-reflective affective-evaluative experience, since it is precisely the content of that experience that now comes under reflective scrutiny. Hence, there is the possibility of what is pre-reflectively experienced as *prima facie* appropriate turning out to be reflectively thought as inappropriate.

So, we might frame the standard of appropriateness as follows:

Standard of appropriateness: do I reflectively judge my *prima facie* appropriate response to be justified in the light of the evaluative feature my emotional response represented its intentional object as having.

Justin D'Arms sums what we are aiming at when assessing our emotions according to this *standard of appropriateness*. He tells us:

Assessments of fittingness are attempts to make sense of or criticise our emotions using standards that speak to the distinctive concerns we take them to embody. It is therefore

important to have a vocabulary that expresses such assessments in particular, as a vehicle for rational interpretation of ourselves and one another. (D'Arms 2005: 10-11)

For example, think of a person who experienced fear of a dog and cites its long hair as a reason for his having been so afraid, having perceived it as dangerous in the light of its long hair. Most people would want to say that the person's affective response in such a case should be judged to be inappropriate, and the perceptual view of the emotions offers a plausible philosophical explanation of this verdict. It says that the emotional response was inappropriate insofar as the reasons sighted are not qualities indicative of the evaluative feature my emotion represented the dog as having, i.e., its being dangerous (the 'distinctive concern' embodied in the emotion of fear). So, given long hair is, in the course of the reflective assessment, accepted by the subject as *not* being a quality indicative of danger, I can come to understand how my emotional response was inappropriate. In this sense what is experienced as *prima facie* appropriate in pre-reflective emotional experience can come to be reflectively judged as inappropriate.

This *standard of appropriateness* is important because it seems to underwrite further intuitions as to what counts as being a relatively well-adjusted subject in terms of one's emotional responses and one's reflections about them. Such that even if our emotional responses and the reasons expressed in them do not always withstand reflective scrutiny (experienced *prima facie* appropriateness by no means guaranteeing judged appropriateness) presumably the possibility, for example, of recognizing by one's own lights what are inappropriate reasons for fear allows one to understand what would be more or less appropriate reasons for fear. Moreover, and relatedly, it holds open the

positive possibility that what are pre-reflectively experienced as appropriate reasons can be reflectively judged or confirmed as appropriate.⁶⁶

Building on these reflections we should note there is a contrast between such a ‘reasons-responsive’ emotional subject (those who in Nietzsche’s terms might be capable of a certain ‘mastering of emotions’ GM II 3), and three other types of emotional subjects who in different ways have a more problematic relation to such standards of appropriateness. We might consider (a) subjects who are entirely alienated from (at least parts of) their emotional lives insofar as they experience them as disclosing no reasons at all. Such a subjects emotional responses are simply rationally unintelligible to them e.g., they can cite no reason for their anger. For these subjects such standards of appropriateness are just not applicable. More complex cases are compulsive subjects whose emotional responses *are* experienced pre-reflectively as caused and merited by their evaluative-features of their objects but the subjects *either*:

(b) *habitually* experience what they themselves reflectively judge to be inappropriate ‘reasons’ for their affective responses in a way not amenable to conscious correction (in this category we find typical cases of emotional recalcitrance, e.g., always perceiving danger on the basis of long hair).⁶⁷

(c) are unable or unwilling to recognize their responses as inappropriate on the basis of the putative reasons they give (e.g., insisting that long hair is an appropriate reason to perceive danger).

⁶⁶ Adam Pelsier makes a similar point although using a slightly different vocabulary: ‘the mere fact that *prima facie* justification conferred by emotional experience is often defeated, however, does not entail that the *prima facie* justification provided by emotion never results in *ultima facie* justification’ (Pelsier 2014: 115, cf. McDowell 1998: 222).

⁶⁷ For further discussion of emotional recalcitrance and its relation to the perceptual theory of the emotions see Tappolet 2012 and Döring 2014.

Presumably case (a) is rare and there is legitimate question as to whether such a subject is really experiencing emotions in any sense that we understand them (perhaps certain psychotic moods are like this). In both the latter cases we are usually in the business, with varying degrees of self-knowledge, of emotional pathologies whereby the standards of appropriateness are in the case of (b) available but do not play their usual function of correcting emotional responses whereby we usually try to, as D'Arms puts it, 'take such steps as we can to feel emotions in accordance with our conclusions';⁶⁸ or in the case of (c) have become ossified in such a way that the individual's own reflective judgements of certain affective responses being appropriate have themselves become habitual and resistant to critical questioning.

However this last case (c) is doubly problematic because we might want to make a distinction between (1) persons who are, perhaps due to forms of psychosis, genuinely unable to recognize the putative reasons they give for their responses as inappropriate, for example they cannot be made to see by any means why citing long hair might be an inappropriate reason to perceive danger, and (2) individuals who for external reasons have cultivated a kind of reflective obstinacy, perhaps through self-deception, regarding their judgements about certain emotional responses, such that they refuse to acknowledge that their emotional response is inappropriate on the basis of the reasons they give because something else of importance is at stake for them in judging a particular emotional response to be appropriate. For example, we might imagine an individual whose jealousy *needs* to be judged to be an appropriate emotional response, reposing on 'good reasons', because to question their *prima facie* appropriate reasons for responding might stand to bring to conscious attention a disturbing fact about their

⁶⁸ D'Arms 2005: 6. Döring seems right to argue that improving one's 'emotional systems' (as she call them) in line with 'better judgements' is not a categorical requirement, it is something we might (if we can) do, but not something that we *ought* to do (Döring 2014: 127).

childhood (i.e., that their habitual jealousy is connected with childhood abandonment issues). In this sense we might say the ‘cognitive cost’ of questioning the standing of their reasons for responding in the way they do, where there might be a path leading from judged inappropriateness to conscious attention of a perturbing fact about their childhood, is too high.

Nevertheless, complex cases aside, the perceptual theory of the emotions seems to have the resources to explain how I can rationally assess some of my emotional responses in terms of whether or not their *prima facie appropriateness* falls short or lives up to being assessed (judged) as *appropriate*, where what we are interested in is whether those reasons withstand reflective scrutiny.

Second, an emotion according to the affective-perception model can be thought to have the possibility of being veridical. We might frame this standard as follows:

Standard of epistemic correctness: as a matter of fact about the world, the object of my emotional response instantiates the evaluative characteristic my emotion represented it as having.

Satisfying such an epistemic standard might in fact be very difficult in certain instances. For example, is there really an evaluative fact of the matter about the Mona Lisa exemplifying beauty which would make my emotional response correct in this sense? Difficult cases aside it seems at least in certain instances such an epistemic standard might be able to be satisfied or indeed unsatisfied.

For example, I might respond with fear to a snake because I perceive its marks at indicating that it is poisonous (i.e., dangerous) but as a matter of fact this snake has had its venom drained and is in fact not poisonous (i.e., not dangerous). My emotional response could be reflectively judged to be appropriate (in the sense of satisfying the above *standard of appropriateness*) but it just turns out that the evaluative feature I took

it to have is in fact not present (my ‘reasons’ are defeasible in this sense). Unsurprisingly then, we might think that the epistemic status of our emotions is related to the veridicality of the beliefs I hold about the presence of certain evaluative features. To mark the contrast between the two standards even more clearly we can observe that they come apart in the opposite direction as well. Since it seems I could in fact have what is the epistemically correct emotional response in terms of matching the supposed ‘matter of fact’ about the world but the response itself could nonetheless be assessed as inappropriate (i.e., reposing on what could be judged as *inappropriate reasons*). For example, I might *not* respond with fear to the snake in the above example, and as it turns out it this is the epistemically correct response insofar as it is *in fact* not dangerous. However, say I did not know that the snake’s venom has been drained and the reason I did not respond with fear was because I take all snakes to be harmless. Here we have an example where what we might say is the correct emotional response reposes on inappropriate reasons (the counter possibility to above, where the incorrect emotional response reposes on appropriate reasons).⁶⁹

So, in sum the perceptual theory of the emotions suggests two ways in which there might be standards of assessment of our emotional responses; at least in certain cases emotions might be thought to have standards of appropriateness and standards of

⁶⁹ Linda Zagzebski considers interesting cases where there can be good grounds for not having an appropriate emotion (say fear of poisonous snakes) which do not undermine that emotion being an appropriate response. Slightly amending her example, consider an individual who does not respond with fear to the poisonous snake because he has trained himself to remain calm in the presence of such snakes. He has done so because he believes it will better serve his survival chances if he is able to avoid fear since he will be less prone to sudden movements which the snakes are more likely to respond aggressively too. Yet even if *he* has good reasons for not responding with fear in such cases, this does *not* show that fear is inappropriate in such cases (see Zagzebski 2014: 175).

epistemic correctness.⁷⁰ I will return to a specific class of veridical emotions that I will argue are central to the FP as Nietzsche's meta-ethical ideal in Section 4 after briefly considering a further objection to the affective-perception view.

3.4 Further Objections and Responses

There exist a number of further objections to the perceptual theory of the emotions, some of which recently have been put forward by Deonna and Teroni.⁷¹ I do not have the space to consider all of them, however there is one in particular they highlight which I want to respond to since they take it show a central disanalogy between sense perception and affective perception of values in a way that might render the theory implausible, and so unable to play the role I have given it thus far in Nietzsche's theory of value and what I will say later about his meta-ethical ideal (see Section 3.5 and Section 4 below).

Deonna and Teroni question what, if emotions are perceptions of value, we are to make of the fact that they seem to non-contingently depend on their non-evaluative 'cognitive bases'.⁷² Taking the example of the public execution, we might say my experience of disgust at perceiving the barbarity of the execution is causally dependent on a sensory complex of seeing, hearing etc. that is ostensibly non-evaluative. So, the objection goes, emotions are 'not independent ways of accessing the objects that exemplify these properties',⁷³ but rather are dependent on their cognitive bases without which they would not be possible, i.e., I have to *see* the execution (sense perception) to

⁷⁰ Todd frames this view in terms of emotional responses being constrained from two sides 'the nature of the response' and 'the world' (Todd 2014: 95).

⁷¹ Deonna and Teroni 2012. See also Brady 2012 and Todd 2014.

⁷² Deonna and Teroni 2012: 69.

⁷³ Ibid: 69.

then be disgusted by it (affective perception). On the other hand ordinary sense perceptions are plausibly thought of as ‘independent ways of accessing the objects’ that they are about.⁷⁴ Therefore ordinary sense perceptions do not require a cognitive base but are the cognitive bases of, amongst other things, emotions (beliefs etc.), hence the disanalogy.

The affective-perception view could respond by arguing that the evaluative properties that the affective-perceptions pick out are higher-order properties that supervene on these non-evaluative natural properties (their ‘cognitive bases’). A good example might be when I see a painting and feel the emotion of admiration on the basis of certain evaluative properties which my affective-perception picks out. Clearly in such instances my affective-perception of the painting as admirable is dependent on the presence of quite specific subtending natural properties of the painting, amongst other things, the finer details of the brushwork, the spatial representation of its features, the specific shades and configurations of colour etc. In this sense we might think that if the ‘natural’ features were changed in some way, say the shades of colour were altered, the painting might well lose its ‘emergent’ higher-order value-qualities. Or consider my response to a piece of Wagnerian music, intended to convey tragic grief, where at precisely the moment where the music score indicates a *crescendo* the orchestra has been directed on this occasion to perform a *calmando*. It is plausible to think that altering the natural features, in this case altering the sonic quality of the orchestra at a specific moment, might alter the ‘perception’ of value in the piece of music.⁷⁵

⁷⁴ Ibid: 69.

⁷⁵ The idea that the evaluative properties are ‘higher-order’ properties that supervene on more basic ones might also provide a response to a different objection, also made by Deonna and Teroni (see their 2012: 68) that there is no ‘sense organ’ of emotion like there is for ordinary sense perception. Since many higher-order properties, like for example expecting a melodic resolution in a piece of music, whilst necessarily dependent on having a certain cognitive base

So it seems the supervenience relations between natural and evaluative features might in certain cases be quite complex. Yet, we should keep in mind that in such instances we need not think that I have to possess detailed thematic awareness of the ‘cognitive bases’ on which my perception of value supervenes. That is to say, it seems plausible to think that my emotion of admiration at the paintings perceived value can be a meaningful response to that evaluative characteristic of the painting, its being admirable, *without* explicit awareness of these ostensibly non-evaluative subtending properties.⁷⁶ So we might conclude that whilst emotions are dependent on their cognitive bases in the way described above, this does not impugn describing at least certain emotions as kinds of higher-order perceptions, namely perceptions of value. Naturally they will be more complex than ordinary sense perception but nonetheless we can still claim that certain emotions are most plausibly described as distinctive kinds of perception. In this sense it is worth bearing in mind that to think of emotions as involving a *kind* of perception (i.e., perception of value) need not involve thinking of them as identical to sensory perceptions in all respects.⁷⁷

(say hearing), do not have a particular sense organ dedicated to picking up for on such melodic ‘requirements’.

⁷⁶ I will leave aside the complex issue of whether emotional experience involves either in part, or exclusively, non-conceptual content (although see fn.33). Although it is worth bearing in mind that insofar as many philosophers have thought, due to ‘fineness of grain’ arguments, that the contents of perceptual experience often outstrip that which is linguistically codifiable in concepts, that if it turned out that emotional experience involved some kind of non-conceptual content in an analogous way this would be a point in favour of the analogy between perceptual experience and emotional experience, and so a point in favour of the perceptual theory of the emotions. For further discussion see Tappolet 2012 and Döring 2014.

⁷⁷ Döring makes an interesting point in claiming that ‘the role of such descriptive features is only *co-constitutive*: value also depends on what we care about’ (Döring 2014: 130). That is to say if I do not care about works of art or aesthetic contemplation a particular painting is hardly likely to show up in my emotional experience as beautiful. Whatever exactly it involved in coming to apprehend a painting or piece of music as sublime, we might think that the meaning of that evaluative predicate is unlikely to be able to captured solely in terms of a complex

However, we might worry about those cases in which the ‘cognitive base’ of one’s emotion is ostensibly non-perceptual. For example, say I hear a news report on the radio describing the conditions at a refugee camp, e.g., the size of the camp, the number of people there, the lack of clean water and food etc. In so hearing the news report I might feel empathy for those who I believe to be suffering, yet we might wonder where the ‘higher-order’ perception of value is in such instances, or put another way, we might ask what perceptual contact I have with the putatively instantiated value property?

What might be said of such cases is that for the individual with the relevant concerns, the one who disvalues human suffering, the evaluative content of the emotion is revealed on the basis of an ‘imagistic’ quasi-perception, which is brought to conscious attention by the ‘natural’ description. Something similar could perhaps be said for those cases where an emotion is caused by something in the past. For example, say someone finds out that they were denied a certain job on spurious grounds many years after the interview took place. Again we might say that part of their emotional experience involves a kind of imagistic remembering of aspects of the interview (the people conducting it, the room etc.), which could be construed in a quasi-perceptual way as involving a sense of the unfairness (disvalue) of the state of affairs that their emotion is directed towards. However, these are merely suggestive responses that would need to be developed further. The more limited aim of this sub-section has been to show that (1) the perceptual model of the emotions is defensible and (2) the analogy with perception is illuminating in crucial respects.

conjunction of natural ‘base’ properties, rather as Döring puts it, our ascription of value in such instances might also ‘depend on our conception of the good life’ (ibid).

3.5 Conclusion to Section: Lawfulness in affective-evaluative experience

Drawing some conclusions with regard to the main themes of this section, we have seen how the account given thus far elucidates the relation between affects and values, underwriting a kind of object-intentionality and normativity. At least some of my affective responses should not be thought of as subjective dispositions or contingent preferences but are experienced as *prima facie* appropriate to the evaluative features of their intentional objects. As such the affective-perception view of values satisfies a central desideratum that I have stressed in this chapter, that of giving us an account of that feature of constraint that is central Nietzsche's account of creative-evaluative experience.

Moreover, in doing so it might also be able to make more sense of the idea of 'lawfulness' that Nietzsche repeatedly stresses in his descriptions of the FP (BGE 213, 262). Since in recognizing values as providing *prima facie* appropriate reasons in this way we might be thought to experience a kind of normative 'lawfulness' insofar as we intuitively grasp what would amount to responding and acting in accordance with those reasons or infringing against them. So, although Nietzsche's talk of 'giving oneself laws' (GS 335) might suggest the need for some conceptually articulated law to follow, in fact some of what he says on this topic implies that what is meant by this idea of 'lawfulness' is the 'law-like' character of certain kinds of (typically pre-reflective) affective-evaluative experiences. Again, his appeal to the 'artist', who we are told has a 'better sense of smell in this matter' (BGE 213) seems apt in this context; 'he knows how strictly and subtly he obeys thousands of laws at this very moment, laws that defy conceptual formulation precisely because of their hardness and determinateness' (BGE 188).

Furthermore, I think a key part of what defines the meta-ethical ideal that the FP represents, and what certainly seems a central dimension to legislation as a metaphor for their evaluative sensibility, might be summed as follows. As we saw above what Nietzsche's meta-ethical ideal involves is an understanding that at least some of our own affective experiences, and the evaluative features of the intentional objects we encounter in them, provide us with law-like (or in my terms *prima facie* appropriate) reasons which 'constrain' us at the pre-reflective level. Yet, the FP would no longer harbor the 'delusion of the contemplative ones' described in GS 301 who 'thinks himself placed as *spectator* and *listener* before the great visual and acoustic play that is life' (GS 301), i.e., he would *not* subscribe to a naïve realism about values which understands them as having the existential status of primary qualities of objects, in the sense of being entirely conceptualizable independent of human sensibility. Rather the FP would have *also* attained an understanding of himself as, in quite a specific sense, the 'giver and granter' of value which constraints him, 'the fact that he is also the actual poet and ongoing author of life' (GS 301). As I explained in Section 3.2, this 'creative' aspect might be understood in terms of the way in which human beings affective experiences constitute the domain of value (*the constitution of value thesis*), and perhaps it is this understanding that Nietzsche is referring to when he says that the FP's "“knowing” is creating, their creating is a legislating” (BGE 211).

So if the account of Nietzsche's meta-ethics above is correct, we might be able to solve Gardner's puzzle (Section 2.1). In recognizing the role of affective states in the 'legislative' mode of evaluation we would not be appealing to the 'causal effects of pre-normative psychological forces',⁷⁸ since as has been argued affects themselves contain what I described as a primitive normativity. As such we would stand to experience such affects as *both* normativity constraining, *prima facie* appropriate to their evaluative-

⁷⁸ Gardner 2009: 16.

objects, *and* as our own; the domain of value being constituted by human affectivity in the sense that what I value is existentially dependent on my affective experiences. In this way affective-experience secures, rather than undermines, understanding myself, in Gardner's terms, as the '*ground* of the values' I affirm.⁷⁹ This seems to be at least part of what Poellner means he claims that, 'affects experienced as one's own, as expressing what one is, would on this construal be those one understands, often-pre reflectively, as being *appropriate* [I would say *prima facie appropriate*], as answerable to and thus as actively relating oneself to the world'.⁸⁰

4. The 'Noble' Mode of Valuing

4.1 Affective Responses to Emotions

The account thus far of the relation between affects and values certainly underwrites an important kind of objectivity and normativity. Yet even if we grant this thesis, when it comes to values we might want something more than this, namely that the evaluative features we pick out in such affective-experiences are, at least some of the time, correct. In this section I want to build on the previous account by explaining how, in favourable circumstances, affective responses to evaluative features of a certain kind of intentional object can be correct, expressing something that could make a claim to being an evaluative fact, and so satisfy what I called *the standard of epistemic correctness* (Section 3.3). Moreover in doing so I want to argue that it is part of what the FP exemplifies as representing Nietzsche's meta-ethical ideal, alongside the self-

⁷⁹ Ibid: 8.

⁸⁰ Poellner 2009: 164.

understanding discussed at the end of the previous Section (3.5), to aim for veridical emotional responses in this way.

So, we might ask, can the claim that, at least in certain cases, a certain emotional response to an object can be known get the evaluative facts *wrong*, and that a different emotional response to the same object can be known to get the evaluative facts right, be defended? Put another way, we might say we are asking about whether emotions can lead us to correct conclusions about evaluative matters of fact. If true this claim might suggest a stronger sense of the objectivity of values by way of a fact of the matter about the evaluative status of at least certain objects of our emotions. However, a problem might arise at this stage with the suggestion of evaluative facts of the matter, namely that this sounds like it implies a stronger form of *metaphysical* realism about values, one that as we noted in Section 3 Nietzsche is opposed to. We might wonder then would such a suggestion be open to J. L. Mackie's argument from queerness which draws attention to way in which so-called evaluative facts of the matter would have to repose on *metaphysically* objective values as 'entities or qualities or relations of a very strange sort, utterly different from anything else in the universe'.⁸¹ In other words, what could stand to be the truth-maker for such evaluative facts of the matter that our emotional responses supposedly pick out, other than some *metaphysically real* value property?

In at least one case that Nietzsche is concerned with we might be able to defend the view that it is possible that *correct* emotions can provide us with evaluative facts of the matter *without* committing to a stronger form of *metaphysical realism* about values (as opposed to the modest *phenomenal* realism about values which I argued is at the centre of Nietzsche's meta-ethics in Section 3). It has been argued that Nietzsche often thinks

⁸¹ Mackie 1977: 38. Although as seen in Section 3.4 one might want to claim that 'higher-order' evaluative properties in some way supervene on lower-order natural properties (even if I as argued there this might not typically be revealed at the pre-reflective level of affective-evaluative experience).

of our emotions as veridical in this way when their intentional objects are *other* affective-evaluative states.⁸² We can think of such cases as second-order emotions, in the sense that they are affective responses to either one's own first-order emotional responses or those of another person. So the main claim of this view is that our second-order affective responses in such cases, the attractiveness or unattractiveness we experience when re-presenting those first-order emotions to ourselves, can reveal what Poellner calls the 'intrinsic phenomenal (dis)value of that state'.⁸³ How though, we might wonder, could such a response be said to pick out an evaluative fact of the matter, and do so in a way somehow less problematically than in cases of where the object of the response is more typical?

In fact the potential success of this view lies in its claim about what stands to be the truth-maker, that is what stands to make true an evaluative fact of the matter, in cases of affective responses to other conscious-affective states. The suggestion is that we might reasonably think that the only thing that can plausibly be appealed to in order to arrive at an evaluative fact of the matter about a conscious-affective state just is its first-personal, experienced attractiveness or unattractiveness as self-transparently re-presented to a conscious subject. The idea is that if we can re-present that first-order affective-state accurately to ourselves then the second-order affective response we experience in doing so provides us with a form of non-inferential evaluative knowledge about that state. There are a series of complex points so an example will be helpful to draw out exactly what this view implies.

Consider the state of mind known as petty egoism that has a characteristic tendency to distort the evaluative features of its objects for self-serving motives. Typically the petty egoist becomes angry in the presence of an individual who appears to represent a challenge to their ego, say for example by having a more extensive knowledge than her

⁸² Poellner 2007: 247.

⁸³ Ibid: 247.

on a particular matter or possessing some ability she does not.⁸⁴ In such a situation the petty egoist will typically provide a justification to herself, and others if necessary, of her anger or frustration, perhaps claiming that the affronting individual is defective in some way, e.g., posing as knowing something of interest when she does not. So we might say that the subject's petty egoism involves, *inter alia*, the characteristic affective response of anger when an affront to the ego is made consciously present, which often leads to a deliberate misrepresentation of the evaluative features of the intentional-objects of that characteristic affective response.

If we plug this example into the view suggested above we come to the conclusion that if the petty egoist (or indeed ourselves) were to self-transparently, that is without deliberate secondary self-deception, re-present that petty egoism, along with the evaluative content of the intentional objects is it distorting (e.g., the impressive knowledge of the affronting individual), she would experience the disvalue of this state in terms of the intrinsic unattractiveness of being in it, that is the experienced disvalue of being affected by a distorting jealousy and anger. That is to say, just as one typically recoils when one realizes that one was motivated by jealousy, the petty egoist would, as now the subject of this accurately re-presented second-order emotional response to the petty egoism, be repulsed by it. In this way that second-order affective response, the experienced unattractiveness of the petty egoism as re-presented, would in fact accurately give us the phenomenal disvalue of this state. Such that we could say that repulsion is the *correct* affective response, and that it provides us with non-inferential (i.e., directly intuited) evaluative knowledge in terms of an evaluative fact of the matter about this kind of petty egoism. It is presumably this phenomenological, evaluative fact of the matter that leads many people to (a) take steps to try to overcome emotions which involve self-deception and distortion like jealousy (one does not want to be in them) (b)

⁸⁴ Note that the affronting individual need not be actually physical present but could have been mentioned in conversation.

avoid accurately re-presenting these aspects of their emotional lives to themselves, or (c) engage in secondary self-deception or denial that one is the bearer of such a state of mind at all, e.g., ““I’m not a jealous person” or “that wasn’t jealousy”.

However, this example highlights the difficulty of the task faced by those trying to reach the correct evaluative conclusion by way of an affective response to such first-order emotions. What is perhaps most difficult, but clearly required if we are to take our responses in such cases as aiming to be correct, is an accurate re-presentation of the intentional objects, along with their evaluative content, of those first-order affective-evaluative states (e.g., the evaluative knowledge of the affronting individual in our example). So we might think there is much room for secondary misrepresentation in terms of getting various features of the first-order emotional experience plainly wrong. Yet if this is the case we should also hold that there is the possibility of getting these features more or less right, that is, with varying degrees of accuracy, successfully re-presenting an emotion and its intentional contents to oneself and by doing so affectively acquainting oneself with an evaluative fact of the matter about it.

Note that the view being put forward here, at least as I am interpreting it, is of a deliberately limited scope. It is *not* claiming that affective responses to evaluative features of intentional objects can be veridical *whatsoever*. Such a claim would (a) require a much more detailed argument, (b) raise the spectre of whether or not this entailed a commitment to a stronger form of *metaphysical* realism about values, and (c) ultimately take us too far afield from Nietzsche’s relation to these issues. Rather the claim is that in a particular subset of cases, namely affective responses that have as their intentional objects other affective-evaluative states, a kind of phenomenological, that is non-inferential and directly experienced, evaluative knowledge can *sometimes* be attained, *just in case* the kinds of conditions outlined in the example (and its explication above) are fulfilled. So the claim is that in such cases we have a good candidate for

something that can satisfy the *standard of epistemic correctness* on emotional responses (see Section 3.4).

By limiting the scope of the claims being made in this way the view need not take a decisive stand of whether or not there might be all kinds of affective-evaluative states that, for whatever reason, perhaps due to the complexity or vagueness of their intentional objects and the values apprehended in them, are *not* able to be re-presented as objects for second-order affective responses in this way. In a similar manner such a view could be agnostic about the controversial issue of whether emotional responses to aesthetic objects like art-works or pieces of music can be correct or incorrect in anything like this way. However, it is worth noting that if certain aesthetic objects were created with the intention of conveying a certain kind of mindfulness in terms of an affective-evaluative attitude, like, for example, the Edward Munch painting *Despair*, then the issue might be more complex and might potentially lend itself to some such account.

4.2 The 'Noble' Mode of evaluation and the Future Philosopher

So, what evidence is there that something like the above view is Nietzsche's, and moreover not only his, but as was claimed at the start of this section a central part of what the FP exemplifies as representing Nietzsche's meta-ethical ideal? In fact we will see below that this way of evaluating, in which affective responses are taken to provide us with a kind of intuitive (or non-inferential) evaluative knowledge of other conscious-affective states, is a key part of what Nietzsche means when he uses the term 'noble'

and, as I will argue below, is central to understanding what the FP exemplifies in representing his meta-ethical ideal.⁸⁵

We can see this mode of affective-evaluation clearly at work, along with its complex mechanics, in BGE 263:

There is an *instinct for rank*, that, more than anything else, is itself the sign of *high* rank; there is a *pleasure* in nuances of respect that indicates a noble origin and noble habits. The subtlety, quality, and stature of a soul is put dangerously to the test when something of the first rank passes by...Anyone whose task and exercise is the investigation of souls will use this very art, in a variety of forms, to establish the ultimate value of a soul, the unalterable, inborn order of rank it belongs to: this sort of investigator will test out the soul's instinct. (BGE 263, cf. BGE 260, 265, 163)

I will argue that this passage once properly explicated (along with some parts of it I have not quoted just yet for simplicities sake) shows the 'noble' mode of evaluation described in 4.1 at work. Firstly, we need to understand that the first sentence is a description of the mode of evaluation Nietzsche is valorising which includes a claim about how to identify it. To make matters more clear I think we should re-phrase the first part of this sentence as follows:

(BGE 263 amended): there is an attunement to value that is itself the sign of an individual being of high value.

⁸⁵ Nietzsche might mean a host of other things when he uses the term 'noble'. Although since many of the sections which exhibit this mode of evaluation are to be found in the final section of BGE, aptly named 'what is noble', I think it fair to say that this mode of evaluation is a central aspect of what he means by his use of this term.

Or in other words there is a *way* of evaluating which when we see it as work we know we are dealing with someone who is exemplary of Nietzsche's meta-ethical ideal. What follows in the passage is a description of an individual whose 'soul's instinct' is being tested to see if he displays such attunement to value by way of his *pre-reflective*, 'before the shudders of authority are there to protect it from intrusive clutches and crudeness' (BGE 263), affective-response when presented with an object of high value, 'when something of the first rank passes by' (BGE 263).

However, up to this point we do not quite have the account I described in 4.1, that of a second-order affective response to first-order emotions or states of mind. In order to see that this is what Nietzsche is gesturing towards we need to specify in more detail what this 'something of the first rank' is that 'passes by' our individual who is being tested. And Nietzsche, albeit quite vaguely, tells us that it is 'something that goes on its way like a *living* touchstone, undiscovered, unmarked, and experimenting, perhaps voluntarily covered and disguised' (BGE 263, my emphasis). I think we should take this to mean that the 'living' something is in fact the 'noble soul', that is the character Nietzsche describes in strikingly similar terms only a couple of sections later as having 'a fundamental certainty...about itself, something that cannot be looked for, cannot be found, and perhaps cannot be lost either. *The noble soul has reverence for itself*' (BGE 287). So the first-order state of mind that our individual who is being tested in BGE 263 is perceptually confronted with is that of *self-reverence*, exemplified by the noble soul, who 'passes by' as a 'living touchstone' for this affective-evaluative state of mind.

Now that the various pieces are in place we are ready to see how our individual fares, that is whether he exemplifies an 'instinct for rank' in terms of a pre-reflective attunement to value by way of the right affective response to the mindful intentional object (the noble's self-reverence). In fact the individual is shown to pass the test with flying colours insofar as his second-order affective response hits the mark. His response

is described by Nietzsche as involving ‘an involuntarily hush, a hesitation of the eye and a quieting of every gesture, all of which indicate that the soul [our individual being tested] *feels* the presence of something deserving the highest honours’ (BGE 263). In other words our individual passes Nietzsche’s test insofar as he displays this requisite ability (attunement to value), shown to us through his affective response of *respect* which ‘the soul [he] *feels*’ to an accurately represented mindful, and in this case noble, intentional object that he is confronted with.

So, what our individual apprehends through this re-presentation of noble-reverence is the positive value of this state of mind, an attunement we might say, to the intrinsic phenomenal value, and so a non-inferential evaluative fact of the matter, about those ‘proud states of soul that are perceived as distinctive’ (BGE 260). He is shown to exemplify Nietzsche’s ‘instinct for rank that...is the sign of *high* rank’ (BGE 263, cf. BGE 163). The lesson of this example is that we can see that there is something highly exemplary for Nietzsche in a mode of evaluation that determines the value of a conscious affective state, primarily at least, through a pre-reflective affective response to it. So to Nietzsche’s question ‘what is noble’ (BGE Section 9), the answer is, at least in part, this mode of evaluation.

However, what evidence is there that this ‘noble’ mode of evaluation is exemplified by the FP? I think the clearest link is where Nietzsche tells us that such FP’s will hold on to the ‘certainty of value standards’ and that ‘these philosophers of the future will demand critical discipline and every habit that leads to cleanliness and rigor in *matters of the spirit*’ (BGE 210, my emphasis cf. BGE 211). I think this preoccupation of FP’s with ‘matters of the spirit’ clearly points towards this ‘noble’ mode of evaluation, whereby the FP should be thought of as an ‘investigator of souls’ who attempts to *accurately* represent the phenomenal value of those ‘states of the soul’ which his affective experience confronts and acquaints him with. Moreover he should be thought

to do so by re-presenting those states to himself, that is ‘test[ing] out the souls instinct’ (BGE 263) and thereby at least attempting to aim for that non-inferential evaluative knowledge, after all Nietzsche describes the FP’s as ‘*those who attempt*’ (BGE 42).

If this is a key part of the what the FP exemplifies, then it is not surprising to find Nietzsche describing this exemplar as committed to a kind of truthfulness, he asks ‘are they new friends of “truth”, these upcoming philosophers?’ and replies ‘probably since all philosophers have so far loved their truths’ (BGE 43). However, Nietzsche also reminds us in the same passage that what he primarily has in mind with his FP is a meta-ethical ideal in terms of an affective-evaluative sensibility; ‘they will not be dogmatists. It would offend their pride, as well as *their taste*, if their truth were a truth for everyone...“My judgement is my judgement: other people don’t have an obvious right to it too” – perhaps this is what such a philosopher of the future will say’ (BGE 43, my emphasis). So, we might say that insofar as Nietzsche tells us that the FP wants to do away with the ‘bad *taste* of wanting to be in agreement with the majority’ (BGE 43, my emphasis), what is exemplary about this figure is that he rejects as phenomenologically unverified *any* evaluative judgement (at least concerning ‘matters of the spirit’) which he is not pre-reflectively acquainted with by means of his ‘noble’ mode of valuation. That is to say his aim is primarily affective attunement to value as revealed through an accurate re-presentation of what we might call the domain of phenomenal value. It is in this sense that Nietzsche describes the FP as holding on to ‘certainty of value standards’, and rejecting as counterfeit any conception of the good as determined by some standard other than as revealed to him in this ‘noble’ way, as for example the good in terms of some appeal to the majority or to some consensus agreement; ‘how could there ever be a “common good”!’ (BGE 43).

Whilst more could be said about this particular theme it is undoubtedly the case that the FP should be seen as committed to an exemplary mode of valuation that would

strongly militate against self-deceptions and misrepresentations. Such that their commitment to *this kind* of truthfulness is a commitment to trying to accurately represent a domain of affectivity and value in which there is, for Nietzsche, much scope for accuracy and a kind of evaluative fact of the matter to be aimed for. As Johnston notes with a similar mode of evaluation in mind, ‘flowing rightly, or being affectively taken with real goods, is itself a compelling ethical ideal’,⁸⁶ and McDowell, ‘one might ask why a training of the feelings (as long as the notion of feeling is comprehensive enough) cannot *be* the cultivation of an ability – utterly unmysterious just because of its connection with feelings – to spot (if you like) the fitness of things’.⁸⁷ Hopefully this final section has showed, why at least in certain specific cases, Nietzsche’s meta-ethical ideal partly involves the cultivation of such ability, a concern signalled in a late work when he equates a ‘free view of reality’ with a ‘refined sense of tact and taste...as body, as gesture, as instinct, - in a word: as reality’ (A 59).

Concluding Remarks

The aim of this chapter was to provide an account of the FP as representing Nietzsche’s meta-ethical ideal. Although some detailed reconstruction and defence of Nietzsche’s meta-ethics in terms of the affective-perception view was required to get to this point I think the ideal as described in Section’s 3.5 and 4.2 captures what this exemplar represents, both in terms of an understanding of value as based in affectivity, and also a ‘noble’ mode of valuation that utilizes this theory of value. Undoubtedly there are other ways to develop Nietzsche’s often-piecemeal reflections on the nature of value, but the hope is that the picture presented here is both exegetically and philosophically plausible,

⁸⁶ Johnston 2001: 193.

⁸⁷ McDowell 1998: 224.

matching up with the claim in my Introduction that Nietzsche's evaluative ideal's are primarily certain kinds of affective-evaluative sensibility, the 'noble mode of evaluation' being one such sensibility, and one that is central to his philosophy of value.

Conclusion to Thesis: Connections and Further Research

In my Introduction I set out the main aim of my thesis as providing an interpretation of Nietzsche's philosophy of value through consideration of his exemplars, looking at different exemplars that can be found in his writings and examining the ideals they represent. I considered the *free spirit* as the critical ideal, Zarathustra as the ethical ideal, and the Future Philosopher as the meta-ethical ideal. In this conclusion I want to address two related concerns. First, it might be asked what the connection between the exemplars, and the ideals they represent, amounts to other than covering the different aspects of Nietzsche's philosophy of value, such that my three chapters will cohere in a more satisfying way if it can be argued that there are substantive connections between them. Across the thesis I have suggested some of these connections, but now want to discuss specific ones in detail. Second, I want to provide some brief ideas for further research on the basis of the interpretation of Nietzsche offered in this thesis.

1.1 Connections between the exemplars

To begin with, one might consider in more detail the way in which the exemplar of the *free spirit*, as representing the critical ideal, clears the necessary space for the development of a positive philosophy of value. Since it might be thought that the latter project is only possible after one's philosophical perspective is liberated from certain metaphysical and religious beliefs and judgements (e.g., belief in contra-causal free will and deserts responsibility, judgements of inexpressible guilt in relation to a perfect God etc.).

Moreover, if matters of taste as kinds of affective-evaluative sensibility are central to Nietzsche's ideal, then we can further see the importance of the critical dimension to his

project. Since the development of an ideal Nietzschean taste, which at this stage could be taken to be comprised of (a) metaphysical indifference (part of the critical ideal), (b) ethical self-overcoming being the foremost practical commitment (the ethical ideal), and (c) an understanding of value as constituted by affective-experiences and an attunement to value in terms of affective disclosures (the meta-ethical ideal), would not be possible if one's taste were distorted by religious and metaphysical feelings and evaluations. That is, if one's affective-evaluations were expressive of what Nietzsche would consider a 'bad' or 'corrupted' taste (e.g., feeling one's existence to be sinful, seeking a state of final eudaimonia in some projected after-life, evaluating the phenomenal *life-world* as of intrinsically less value than a speculative metaphysical realm etc.) In fact Nietzsche is explicit about how what he calls the 'severe suspicion', which I think we should interpret as referring to the critical role of the *free spirit*, allows for the development of an improved 'second taste':

One thing is strangest of all: after all this, one will have a different taste - a second taste. From such abysses, also from the abyss of severe suspicion, one returns newborn, having shed one's skin...with a more tender tongue for all good things, with merrier senses, with a more dangerous second innocence in joy, more childlike, and at the same time a hundred times subtler than one had ever been before. (NCW E2, cf. BGE 224, A 59)¹

The second connection between the exemplars I want to draw out, in relation to the *free spirit*, is the how the evaluative attitude of metaphysical indifference, which it was argued *free spirit* exemplifies, relates to questions concerning the metaphysical status of the values involved in Nietzsche's project. Since, if Nietzsche's meta-ethics is

¹ This passage is an amended version of section 4 of the preface to GS.

interpreted as a modest, or unmetaphysically demanding *phenomenal* realism,² then there remains a further question about metaphysical status of the values that are revealed at this phenomenal level (i.e., in affective-evaluative experience). Such inquiry has often taken the form of asking, from one side of the debate, whether values are merely subjective or ‘projected’ Lockean secondary qualities. And when such questioning is answered in the affirmative this empiricist error theory about the phenomenology of value is offered alongside an anti-realist metaphysics of value, which claims that values do not figure in our most fundamental descriptions of reality.³

On the other side it has often been claimed that phenomenal value experience enjoys correspondence, of some kind or gradation, with a postulated metaphysical value-object. From a Platonistic perspective, the suggestion is that much of our phenomenal value-experience involves a lower representation of a higher metaphysical value (e.g., the ‘Form of the Good’), whereby the latter is located in a mind-independent reality. Alternatively, whilst remaining on the other side to anti-realist positions, one could opt for a primary quality model of values, as a kind of *metaphysical* value-realism, whereby value experience discloses a quality of an object whose ontological reality is independent of human experience. On such a view the value qualities of objects would have the same metaphysical status that is typically accorded to primary qualities (at least on a Lockean reading) of objects, such as form, extension etc., and so would *not* be essentially phenomenological in character.

I want to argue that Nietzsche’s ideal, who possesses a higher taste involving the evaluative attitude of metaphysical indifference, would have reasonable grounds for being unmoved by such speculations about the metaphysics of value (I shall say more on what basis below). More specifically the suggestion is that such metaphysical speculation should not undermine projects that depend on value being ‘real’ (, in the

² See Chapter 3, Section 3. Poellner 2007: 227-261.

³ See Mackie 1977.

sense that was given an account of in Chapter 3), like for example an ethics of self-overcoming.

Building up to this position, it should be kept in mind that for Nietzsche, at least according to the interpretation given of his meta-ethics in Chapter 3, value is primarily disclosed through the registering of value in one's affective responses to the intentional objects of those experiences. Questions about the metaphysics of value that abstract from how value is revealed in this way, such that this phenomenology of value is either ignored, or taken to be an 'appearance' which is in systematic 'error' according to some metaphysically-loaded alternative theory, would be fair candidates for the attitude of indifference, or so I will suggest.

What views might fall under the scope of this attitude in this context? We might think, for example, that the metaphysical dimension to projectivist meta-ethics, according to which there exists a value-free reality onto which we project our internal Humean sentiments, plausibly would. As would the Platonic view of a metaphysically mind-independent reality in which truly objective values reside, like the aforementioned 'Form of the Good', which are not dependent, for their being, on human affective-experiences.⁴ However, as I questioned in Chapter 1 with regard to metaphysical indifference concerning 'Thin Metaphysical' truths,⁵ the grounds for a dismissal of this form of inquiry, and any of its putative results, needs to be accounted for in more detail. Below I want to transpose what I argued was the most plausible way of interpreting that attitude onto questions of value.

I argued in Chapter 1 that metaphysical indifference can be read in terms of skepticism about whether answers to the fundamental questions that occupy TM ('thin

⁴ See McDowell 1998: 224-226 for further discussion of the relation between the way value is revealed at the phenomenal level and those alternative 'metaphysically loaded' theories of value.

⁵ See Chapter 1, Section 2.

metaphysics') are ever likely to be forthcoming, where what would be required for some such answer to count as knowledge (i.e., *justified*, true belief) is something approaching consensus amongst component inquirers, a goal that has eluded metaphysical philosophy for over 2000 years. Drawing on this, we might think a similar point could be made with regard to questions concerning the metaphysics of value. Since disagreement regarding the latter has been, and continues to be, no less persistent than over more traditional metaphysical questions. So, in a Nietzschean spirit, one might think that what he would likely recommend is not more intricate arguments (or counter arguments) for any one of positions canvassed above concerning the metaphysical status of values, but rather (a) 'genealogical-cum-psychological' diagnoses of the underlying motivations of those philosophers who continue to be engaged by such speculations,⁶ and (b) renewed attention to the phenomenology of value. Since it could be argued that inquiry into the metaphysics of value has often been pursued to the detriment of attention to that which stands to be revealed in a self-transparent, and accurately described, experience of value. That is, one might think, the resources for a proper understanding of the transcendental structure (i.e., condition of possibility) of value as *essentially phenomenological* (cf. GS 301). As Nietzsche remarks, 'would it not be quite probable...that precisely the most superficial and external aspect of existence – what is most apparent; its skin and its sensualization – would be grasped first and might even be the only thing that let itself be grasped (GS 373). This understanding is one which Nietzsche thinks is overlooked by Platonists,

⁶ One such 'psychological' diagnosis is provided by Johnston 2001 in terms his claim to reveal some of the psychological motivations behind certain kinds of projectivism. Nietzsche himself claims to reveal the psychological motivations behind the understanding of value given in Platonism and Christianity by identifying them as forms of the ascetic ideal, whereby they deny the value of the human *life-world*, and in doing so distort the nature of value as existentially dependent on human beings (see BGE 1-2, GM III 11, 24, 25).

Religious Philosophies of various stripes, and, albeit in a different way, and for different reasons, Humean or empiricist-inspired meta-ethics like Mackie's error theory.⁷

Whilst more could be said about this stance of metaphysical indifference transposed on to the evaluative domain, it is interesting to see how the critical ideal, as represented by the exemplar the *free spirit*, has a role to play beyond 'severe suspicion' (NCW E2), insofar as indifference might be one possible stance to take regarding questions concerning the metaphysical status of the values involved in Nietzsche's more positive ethical and meta-ethical ideals.

The next connection I want to draw is between the exemplar of Zarathustra, who represents Nietzsche's ethical ideal of self-overcoming, and the Future Philosopher, who represents both an understanding of the way value is constituted (i.e., by affective experiences) and exemplifies a 'noble' mode of evaluation. One simple link between these exemplars could be explained in terms of the way in which normative ethics is traditionally thought to be dependent on meta-ethics. In this sense whatever evaluative practices are involved in self-overcoming would not have clay-feet, but could be argued to have their grounding in a meta-ethics that articulates a sense in which values are part of the 'fabric of the [human] world' (to use Mackie's phrase in a service of a view different from his own 'error theory').⁸ In this sense instances of self-overcoming would not be expressive of a naivety which, upon philosophical reflection, might be revealed to involve extravagant 'realistic' assumptions (a conclusion that might follow if Nietzsche is interpreted as committed to some kind of 'fictionalism' about *all* value claims, which is perhaps another reason to be skeptical about such readings).⁹ As such, it is important for Nietzsche's ethics of self-overcoming, or any normative recommendations about what kinds of evaluative practices are more worthwhile and

⁷ See Mackie 1997.

⁸ Ibid: 15.

⁹ See Hussain 2007: 157-191 and Reginster 2006: 85.

offers suggestions about the way in which they should be pursued (even if, as in Nietzsche's case, such normative ethics is often offered in a quasi-particularist way), to be supported by a meta-ethics that provides a basis for at least *some* of those values being in some sense 'real', 'objective' and normatively constraining in a way that they have to be if they are to be rationally intelligible to the subjects who make them.¹⁰

With a different emphasis one might ask what the connection is between the 'noble' mode of evaluation, in terms of that attunement to value described at the end of Chapter 3, and the commitment to *continual* self-overcoming that I argued represents the highest demand of Nietzsche's normative ethics. Since, one might think that whatever values are revealed in that 'noble' mode could come to be articulated in a fully conceptualized, self-evaluative framework, and it might be argued that such a 'fixing of value' conflicts with *continual* self-overcoming; that is to a never-ending re-evaluation of our fundamental evaluations which would see any fully conceptualized fundamental evaluations as always falling short of what we were after in striving towards our highest ideals.

Whilst there does seem to be a conflict here I think once we remind ourselves of what exactly is involved in the 'noble' mode of evaluation these two aspects of an ideal Nietzschean taste can be complementary. Nietzsche himself suggests as much when he tells us that 'good taste' demands that one should not 'strip [existence] of its *ambiguous* character', and that 'good taste' is above else 'the taste of reverence for everything that lies beyond your *horizon*' (GS 373, both emphases are mine).

Firstly, we might say that there is no reason to think that the attunement to value characteristic of the 'noble' mode of evaluation is a once and for all matter, i.e., revelation of value on one occasion and then a static fixing of that value. Rather, the disclosure of value in the 'noble' way could involve the development of an affective-

¹⁰ For this argument about the link between values as 'real' and 'objective' and rational intelligibility see Chapter 3, Section 3.

evaluative sensibility that was always working towards clearer and more self-transparent disclosure of the value of the objects it is being confronted with. Such a project might be described in Nietzsche's own terms as:

Collecting material, formulating concepts, and putting into order the tremendous realm of tender value feelings and value distinctions that live, grow, reproduce, and are destroyed, – and, perhaps, attempting to illustrate the recurring and more frequent shapes of this living crystallization. (BGE 186)

Yet, as stated in Chapter 2 in considering the horizontal nature of ethical ideals, the project Nietzsche describes here could be a legitimate one even if it were recognized that final, or once and for all, articulation of the highest values in determinate or fixed value-concepts is neither entirely possible nor satisfactory. It is in this way that Nietzsche speaks of value as something that 'lives' in the sense of its being able to 'grow, reproduce...[and be] destroyed' (BGE 186). Moreover, we might also add that such a project would most likely have to be a continual one because of the ever-present threat of the kinds of self-deceptions and distortions that might skew such 'open' attunement to value (e.g., forms of ethical complacency or motivated reflective distortion of value as exemplified by Nietzsche's *ressentiment subjects* in GM I).

Secondly, we can see how an ethics of self-overcoming and the meta-ethical ideal of attunement to a 'growing' and 'reproducing' good might align by noting that it is our sense of the 'good', as our most fundamental value, that is constitutive, *inter alia*, of those frameworks of self-evaluation that are up for re-evaluation in Nietzschean self-overcoming. In this sense an attunement to value as one that both (a) reveals what it is that we value most highly and (b) that what we value most highly is always developing and open to change and revision, could be thought to be a necessary condition for the kind of *continual* overcoming discussed in Chapter 2.

Finally, it should be kept in mind that the way value is revealed in the ‘noble’ mode of evaluation is primarily at the pre-reflective level of affective experience, described by Nietzsche in terms as an apprehension of value that is ‘colourful, young and malicious, so full of thorns and secret spices’ (BGE 296). That attunement to value and the evaluations it provides us with would not, at least not initially, be experienced by way of determinate or fixed value-concepts, and so perhaps what is revealed therein would not necessarily be authentically *expressed* in a fully articulated, conceptual frameworks of value. To remind the reader of a passage quoted in Chapter 2, Nietzsche says that ‘our true experiences are completely taciturn. They could not be communicated even if they wanted to be. This is because the right words for them do not exist. The things we have words for are also the things we have already left behind’ (TI VIII 26). Naturally, the connection I am suggesting here between pre-reflective affective-experience as a source for an intuitive apprehension (or non-inferential knowledge) of value and that anti-conceptualist strand in Nietzsche thinking about ethical ideals discussed in Chapter 2 would need further development. Although as with the other points above it shows an important connection between the ethical ideal of self-overcoming exemplified by Zarathustra and the meta-ethical ideal of the Future Philosopher.

1.2 Ideas for further research

I now want to briefly suggest some further avenues for research on the basis of the account of Nietzsche’s philosophy of value given in this thesis.

Firstly, I think something more substantial could be developed in terms of a Nietzschean view of practical reason on the basis of the affective-perception model of value given in Chapter 3. Whilst deliberative or reflective reason often seems secondary for Nietzsche, it could be argued that something more constructive could be said in this

area, whilst keeping in mind his often-critical remarks on the role of the ‘will’ and reflective agency. Since the language of ‘commanding’ and ‘willing’ that pervades many of the later descriptions of the FP (BGE 203, 209, 211, 213) suggests that a more positive interpretation of the role of reflective deliberation, particularly in those situations where we experience a conflict of values and commitments, could be developed.

Secondly, in terms of pursuing the methodology of considering Nietzsche’s ideals through the exemplars that represent them, one might look to a number of other exemplary figures. For example, there is ‘sovereign individual’ of GM II, a figure who we are told possesses a kind of ‘protracted will’ which suggests something more positive in terms of a kind of reflective agency.¹¹ One might also look at the more negative exemplars to sharpen the critical focus of Nietzsche project, such as the *men of resentment*, the *ascetic priests* (GM I and III) or even the complex counter-ideal that Wagner is often claimed to be exemplary of. This method is a fruitful one because as so often with Nietzsche his ideals and counter ideals are not represented, or at least not exclusively, in terms of abstract systems of belief or formal positions, but through such exemplars.

Yet, as was stated originally in my Introduction this is no mere idiosyncrasy on Nietzsche’s part, but rather is central to his attempt to (a) pursue a philosophy of value within a critical framework that is skeptical about the rational warrant for externally imposed prescriptions and categorical imperatives, and (b) more accurately account for certain kinds of affective-evaluative sensibilities. This latter aspect of his project therefore benefits from an appeal to a certain kind of ‘lived experience’ that exemplars are more than suitable to fit, allowing Nietzsche to draw out and make vivid that which is of most interest in the both the kinds of evaluative outlooks he is, and is not, in favor

¹¹ See Ridley 2009: 181-196, Gemes 2009: 33-50 and for the critical response Leiter 2011: 101-119.

of. It is therefore my hope that this thesis has shed some light on Nietzsche's philosophy of value understood in these terms, having both given (1) a comprehensive picture of the substance and content of Nietzsche's ideal in terms of his evaluative ideals as represented by his exemplars, and (2) at least some sense of Nietzsche's grounds (or reasons) for his evaluative preferences as given in those evaluative ideals.

Bibliography

Works by Nietzsche

Nietzsche, F. (2005) *The Anti-Christ*. Ed. A. Ridley and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.

Nietzsche, F. (2002) *Beyond Good and Evil*. Ed. R.P. Horstmann and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.

Nietzsche, F. (2005) *The Case of Wagner*. Ed. A. Ridley and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.

Nietzsche, F. (1997) *Daybreak*. Ed. M.Clark and B. Leiter and. Trans. R. J. Hollingdale. Cambridge: Cambridge University Press.

Nietzsche, F. (2005) *Ecce Homo*. Ed. A. Ridley and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.

Nietzsche, F. (2001) *The Gay Science*. Ed. B. Williams. Trans. J. Nauckhoff. Cambridge: Cambridge University Press.

Nietzsche, F. (1997) *On the Genealogy of Morality*. Ed. K.A. Pearson. Trans. C. Diethe. Cambridge: Cambridge University Press.

Nietzsche, F. (2001) *Human, All Too Human* (1). Trans. G. Handwerk. Stanford: Stanford University Press.

Nietzsche, F. (2005) *Twilight of the Idols*. Ed. A. Ridley and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.

Nietzsche, F. (1968) *The Will to Power*. Ed. W, Kaufmann Trans. Walter Kaufmann and R. J. Hollingdale. New York: Vintage Books.

Nietzsche, F. (2006) *Thus Spoke Zarathustra*. Ed. A.D. Caro and R. Pippin. Trans. A D. Caro. Cambridge: Cambridge University Press.

- Nietzsche, F. (1995) *Unfashionable Observations*. Trans. R. T. Gray. Stanford: Stanford University Press.
- Nietzsche, F. (2003) *Writings from the Late Notebooks*, trans. K. Sturge, R. Bittner (ed). Cambridge: Cambridge University Press.
- Nietzsche, F. (2005) *Nietzsche Contra Wagner*. Ed. A. Ridley and J. Norman. Trans. J. Norman. Cambridge: Cambridge University Press.
- Nietzsche, F. (1967-) *Werke: Kritische Gesamtausgabe*. Colli, G. and Montinari, M, et al. (eds.) Berlin: de Gruyter.

Other works

- (2006) *The Bible* (King James Version with The Apocrypha, Norton, D (ed). London: Penguin Books.
- Anderson, R. A. (2011) ‘On the nobility of Nietzsche’s priests’ in in S. May (ed) *Nietzsche’s On the Genealogy of Morality*, 24-55. Cambridge: Cambridge University Press.
- (2012) ‘What is a Nietzschean Self’ in C. Janaway and S. Roberston (eds) *Nietzsche, Naturalism and Normativity*, 202-235, Oxford: Oxford University Press.
- Anselm (1998) *The Major Works*. Oxford. Oxford University Press.
- Aquinas, T. (1993) *Selected Philosophical Writings*. Oxford: Oxford University Press.
- Aristotle (1962) *The Politics*, trans. T.A. Sinclair. Harmondsworth: Penguin.
- Aristotle (1998) *The Nichomachean Ethics*, trans. D. Ross. Oxford: Oxford University Press.
- Brady, M. (2012) ‘Emotions, Perceptions and Reasons’ in C. Bagnoile (ed) *Morality and the Emotions*, 135-49. Oxford: Oxford University Press.
- Clark, M. (1990) *Nietzsche on Truth*. Cambridge: Cambridge University Press.

- (1998) ‘On Knowledge, Truth, and Value...’ in C. Janaway (ed) *Willing and Nothingness: Schopenhauer as Nietzsche’s Educator*, 37-78. Oxford: Clarendon Press.
- Clark, M. and Dudrick, D. (2007) ‘Nietzsche and Moral Objectivity: The Development of Nietzsche’s Metaethics’ in B. Leiter and N. Sinhababu (eds) *Nietzsche and Morality*, 192-226. Oxford: Oxford University Press.
- Cohen, J. (2010) *Science, Culture and Free Spirits*. New York: Humanity Books.
- Dancy, J. (1993) *Moral Reasons*. Oxford: Blackwell.
- (2002) *Practical Reality*. Oxford: Oxford University Press.
- D’Arms, J. (2005) ‘Two Arguments for Sentimentalism’ in *Philosophical Issues* 15 (Normativity), 1-12.
- Davidson, D. (1980) *Essays on Actions and Events*. New York: Oxford University Press.
- Deleuze, G. (1977) ‘Nomad Thought’ in D. B. Allison (ed) *The New Nietzsche*, 142-150. New York: Dell Publishing.
- (1986) *Nietzsche and Philosophy*, trans. H. Tomlinson. London: Continuum.
- Deonna, J. and Teroni, F. (2012) *An Introduction to the Philosophy of the Emotions*. London: Routledge.
- Derrida, J. (1977) ‘The Question of Style’ in D. B. Allison (ed) *The New Nietzsche*, 176-190. New York: Dell Publishing.
- Descartes, R. (1996) *Meditations on First Philosophy*, trans. J. Cottingham, Cambridge: Cambridge University Press.
- (1998) *Discourse on Method*, trans. D. A. Cress, Cambridge: Hackett Publishing.
- Döring, S. (2007) ‘Seeing What to Do: Affective Perception and Rational Motivation’ in *Dialectica*, 363-394.

- (2014) ‘Why Recalcitrant Emotions are Not Irrational’ in S. Roeser and C. Todd (eds) *Emotion and Value*, 124-136. Oxford: Oxford University Press.
- Foucault, M. (2001) ‘Nietzsche, Genealogy, History’ in J. Richardson and B. Leiter (eds) *Nietzsche*, 341-360. Oxford: Oxford University Press.
- Frankfurt, H. (1998) *The importance of what we care about*. Cambridge: Cambridge University Press.
- (1999) *Necessity, Volition, and Love*. Cambridge: Cambridge University Press.
- Gardner, S. (2009) ‘Nietzsche, the Self, and the Disunity of Philosophical Reason’ in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 1-32. Oxford: Oxford University Press.
- Gemes, K, and Janaway, C. (2005) ‘Naturalism and Value in Nietzsche’ in *Philosophy and Phenomenological Research* 71 (3), 729-740.
- Gemes, K. (2009) ‘Nietzsche on Free Will, Autonomy and the Sovereign Individual’ in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 33-50. Oxford: Oxford University Press.
- Han-Pile, B. (2009) ‘Nietzsche and Amor Fati’ in *European Journal of Philosophy* 19 (2), 224-261.
- Harcourt, E. (2011) ‘Nietzsche and the “aesthetics of character”’ in S. May (ed) *Nietzsche’s On the Genealogy of Morality*, 265-285. Cambridge: Cambridge University Press.
- Hare, R. M. (1963) *The Language of Morals*, Oxford: Clarendon Press.
- Havas, R. (2013) ‘The Overman’ in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 461-484. Oxford: Oxford University Press.
- Hegel, G. W. F. (1977) *Phenomenology of Spirit*, trans, J. N. Findlay. Oxford: Oxford University Press.

- Heidegger, M. (2005) *Being and Time*, trans. J. Macquarrie and E. Robinson Oxford: Blackwell Publishing.
- Hempel, C. (1945) 'Studies in the Logic of Confirmation' in *Mind*. 54 (213), 1-26.
- Hume, D. (1978) *A Treatise of Human Nature*. Oxford: Oxford University Press.
- (1987) *An Enquiry Concerning the Principles of Morals*. Cambridge: Hackett Publishing.
- (1998) *Dialogues Concerning Natural Religion*. Cambridge: Hackett Publishing.
- Hurka, T. (2007) 'Nietzsche: Perfectionist' in B. Leiter and N. Sinhababu (eds) *Nietzsche and Morality*, 9-31. Oxford: Oxford University Press.
- Husserl E. (1970) *The Crisis of European Sciences and Transcendental Phenomenology*, trans. D. Carr. Evanston: Northwestern University Press.
- (2005) *Phantasy, Image Consciousness and Memory*, trans. J. B. Brough. Dordrecht: Springer.
- Hussain, N. (2007) 'Honest Illusion: Valuing for Nietzsche's Free Spirits' in in B. Leiter and N. Sinhababu (eds) *Nietzsche and Morality*, 157-192. Oxford: Oxford University Press.
- Janaway, C. (2007) *Beyond Selflessness*. Oxford. Oxford University Press.
- (2009) 'Autonomy, Affect and the Self...' in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 51-68. Oxford: Oxford University Press.
- (2012) 'Nietzsche on Morality, Drives, and Human Greatness' in C. Janaway and S. Roberston (eds) *Nietzsche, Naturalism and Normativity*, 183-201, Oxford: Oxford University Press.
- Jaspers, K. (1997) *Nietzsche: An Introduction to the Understanding of His Philosophical Activity*. Baltimore: John Hopkins University Press.
- Johnston, M. (2001) 'The Authority of Affect' in *Philosophy and Phenomenological Research* 63: 181 – 214.

- Kant, I. (1991) *The Moral Law (Groundwork of the Metaphysic of Morals)*, trans. H. J. Paton, London: Routledge.
- (1998a) *The Critique of Pure Reason*, trans. P. Guyer, Cambridge: Cambridge University Press.
- (1998b) *Religion within the Boundaries of Mere Reason*, trans. A. Wood, Cambridge: Cambridge University Press.
- Katsafanas, P. (2005) 'Nietzsche's Theory of Mind' in *European Journal of Philosophy* 13, 1-33.
- (2011) 'Deriving ethics from action: a Nietzschean version of Constitutivism' in *Philosophy and Phenomenological Research* 83, 620–60.
- (2013a) *Agency and the Foundation of Ethics: Nietzschean Constitutivism*. Oxford: Oxford University Press, 2013.
- (2013b) 'Nietzsche's Philosophical Psychology' in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 727-756. Oxford: Oxford University Press.
- (2015) 'Value, Affect, Drive' in P. Kail and M. Dries (eds) *Nietzsche on Mind and Nature*, 163-188. Oxford: Oxford University Press.
- Kaufmann, W. (1974) *Nietzsche*. Princeton: Princeton University Press.
- Klages, L. (1926) *Die Psychologischen Errungenschaften Nietzsches*. Leipzig: Barth.
- Korsgaard, C. (1996) *The Sources of Normativity*. Cambridge: Cambridge University Press.
- (2009) *Self-Constitution: Agency, Identity and Integrity*. Oxford: Oxford University Press.
- Leiter, B. (1992) 'Nietzsche and Aestheticism' in *Journal of the History of Philosophy*, Vol. 30, No. 2, 275-290.
- (1997) 'Nietzsche and the Morality Critics' in *Ethics*, Vol. 107 (2), 250-285.

- (2000) ‘Nietzsche’s Metaethics: Against the Privileged Readings’ in *European Journal of Philosophy* 8: 277-297.
- (2011) ‘Who is the “sovereign individual”? Nietzsche on freedom’ in S. May (ed) *Nietzsche’s On the Genealogy of Morality a critical guide*, 101-119. Cambridge: Cambridge University Press.
- (2015) *Nietzsche on Morality* (second edition). London: Routledge.
- Leon, M. (1992) ‘Rationalizing Belief’ in *Philosophy Papers* 21 (3), 299-314.
- Lowith, K. (1956) *Nietzsche’s Philosophie der ewigen Wiederkehr des Gleichen*. Stuttgart.
- Mackie, J. L. (1977) *Ethics: Inventing Right and Wrong*. London: Penguin.
- (1982) *The Miracle and Theism*. Oxford: Clarendon Press.
- MacIntyre, A. (1981) *After Virtue*. London: Duckworth.
- Meyer, M. (2004) ‘Human, all too Human and the Socrates Who Plays Music’ in *International Studies in Philosophy*, 36 (3), 171-182.
- McDowell, J. (1985) ‘Values and Secondary Qualities’ in J. Rachels (ed) *Ethical Theory I*, 210-226. Oxford: Oxford University Press.
- (1994) *Mind and World*. Cambridge: Harvard University Press.
- Misak, C. F (1995) *Verificationism Its History and Prospects*. London: Routledge.
- Nagel, T. (1998) ‘Value’ in J. Rachels (ed) *Ethical Theory I*, 109-124. Oxford: Oxford University Press.
- Nehamas, A. (1985) *Nietzsche: Life as Literature*. London: Harvard University Press.
- (2001) ‘How One Becomes What One Is’ in J. Richardson and B. Leiter (eds) *Nietzsche*, 255-280. Oxford: Oxford University Press.
- Nussbaum, C. N. (2001) *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

- Pelser, A. (2014) 'Emotion, Evaluative Perception, and Epistemic Justification' in S. Roeser and C. Todd (eds) *Emotion and Value*, 107-123. Oxford: Oxford University Press.
- Pippin, R. (2006a) 'Agency and Deed in Nietzsche's *Genealogy of Morals*' in K. A. Pearson (ed) *A Companion to Nietzsche*, 371-387. Oxford: Blackwell Publishing.
- (2006b) 'Introduction' in *Thus Spoke Zarathustra*. Ed. A.D. Caro and R. Pippin. Trans. A D. Caro. Cambridge: Cambridge University Press.
- (2009) 'How to Overcome Oneself: Nietzsche on Freedom' in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 69-88. Oxford: Oxford University Press.
- Plato (2008) *The Symposium*. Ed M. C. Howaston and Frisbee C. C. Sheffield. Cambridge: Cambridge University Press.
- Poellner, P. (1995) *Nietzsche and Metaphysics*. Oxford: Clarendon Press.
- (2001) 'Perspectival Truth' in J. Richardson and B. Leiter (eds) *Nietzsche*, 85-117. Oxford: Oxford University Press.
- (2006) 'Phenomenology and Science in Nietzsche', in K. A. Pearson (ed) *A Companion to Nietzsche*, 297-313. Oxford: Blackwell Publishing.
- (2007) 'Affect, Value and Objectivity' in B. Leiter and N. Sinhababu (eds) *Nietzsche and Morality*, 227-261. Oxford: Oxford University Press,
- (2009) 'Nietzschean Freedom' in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 151-180. Oxford: Oxford University Press.
- (2011) 'Ressentiment and Morality' in S. May (ed) *Nietzsche's On the Genealogy of Morality a critical guide*, 120-141. Cambridge: Cambridge University Press.
- (2012) 'Aestheticist Ethics' in C. Janaway and S. Roberston (eds) *Nietzsche, Naturalism and Normativity*, 52-80, Oxford: Oxford University Press.
- (2013) 'Nietzsche's Metaphysical Sketches...' in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 675-700. Oxford: Oxford University Press.

- Raz, J. (2002) *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.
- Rée, P. (2003) *The Origin of the Moral Sensations*, trans. R. Small. Illinois: University of Illinois Press.
- Reginster, B. (2003) 'What is a Free Spirit? Nietzsche on Fanaticism in *Archiv für Geschichte der Philosophie* 85 (1), 51-85.
- (2006) *The Affirmation of Life*. London: Harvard University Press.
- (2011) 'The genealogy of guilt' in S. May (ed) *Nietzsche's On the Genealogy of Morality*, 56-77. Cambridge: Cambridge University Press.
- (2013a) 'The Psychology of Christian Morality' in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 701-727. Oxford: Oxford University Press.
- (2013b) 'Honest and Curiosity in Nietzsche's Free Spirits' in *Journal of the History of Philosophy* 51 (3), 441-463.
- Richardson, J. (2009) 'Nietzsche's Freedoms' in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 127-150. Oxford: Oxford University Press.
- (2013) 'Nietzsche on Life's Ends' in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 756-784. Oxford: Oxford University Press.
- Ridley, A. (2005a) 'Nietzsche on Art and Freedom' in *European Journal of Philosophy* 15 (2), 204-224.
- (2005b) 'Introduction' in *Twilight of the Idols*. Ed. A. Ridley and J. Norman, trans. J. Norman Cambridge: Cambridge University Press.
- (2007) *Nietzsche on Art*. London: Routledge.
- (2009) 'Nietzsche Intentions: What the Sovereign Individual Promises' in K. Gemes and S. May (eds) *Nietzsche on Autonomy and Freedom*, 181-196. Oxford: Oxford University Press.

- Robertson, S. (2012) 'The Scope Problem...' in C. Janaway and S. Robertson (eds) *Nietzsche, Naturalism and Normativity*, 81-110, Oxford: Oxford University Press.
- Rorty, R. (1989) *Contingency, Irony and Solidarity*. Cambridge: Cambridge University Press.
- Sade, M. (1991) *The 120 Days of Sodom*. London: Arena Books.
- Salome, L. (2001) *Nietzsche*. Illinois: University of Illinois Press.
- Sartre, J. P. (1948) *Existentialism and Humanism*, trans. P. Mairet. London: Menthuen.
- (1996) *Notebooks for an Ethics*, trans. D. Pellauer. London: University of Chicago Press.
- (2003) *Being and Nothingness*, trans. H. Barnes London: Routledge.
- Scheler, M. (1974) 'The Meaning of Suffering' in M. S Frings (ed) *Max Scheler (1874-1928) Centennial Essays*. The Hague.
- Schopenhauer, A. (1969a) *The World as Will and Representation Vol. I*, trans. E. F. J. Payne. Dover Edition.
- (1969b) *The World as Will and Representation Vol. II*, trans. E. F. J. Payne. Dover Edition.
- Searle, J. *Intentionality: An Essay in the Philosophy of Mind*. New York: Cambridge University Press, 1983.
- Solomon, R. (1976) *The Passions*. Doubleday.
- Strawson, G. (1994) 'The Impossibility of Moral Responsibility' in *Philosophical Studies* 75, 5-24.
- Strawson, P. F. (2008) *Freedom and Resentment*. Routledge: New York.
- Stroud, B. (1993) 'Gilding and Staining the World with Sentiments and Phantasms' in *Hume Studies*, Vol. XIX (2), 253-272.
- Taylor, C. (1982) 'Responsibility for Self' in G. Watson (ed) *Free Will*, 111-126. Oxford: Oxford University Press.

- (1985) ‘What is Human Agency’ in *Human Agency and Language*. London: Cambridge University Press.
- (1989) *Sources of the Self*. Cambridge: Cambridge University Press.
- Tappolet, C. (2012) ‘Emotions, Perceptions, and Emotional Illusions’ in Calabi, C. (ed) *Perceptual Illusion. Philosophical and Psychological Essays*. New York: Palgrave-Macmillan.
- Tevenar, G. (2013) ‘Zarathustra” ‘That Malicious Dionysian”’ in K. Gemes and J. Richardson (eds) *The Oxford Handbook of Nietzsche*, 272-297. Oxford: Oxford University Press.
- Todd, C. (2014) ‘Relatively Fitting Emotions and Apparently Objective Values’ in S. Roeser and C. Todd (eds) *Emotion and Value*, 90-106. Oxford: Oxford University Press.
- Velleman, D. (2000) *The Possibility of Practical Reason*. Scholarly Publishing Office University of Michigan Library.
- (2006) *Self to Self. Selected Essays*. Cambridge: Cambridge University Press.
- Watson, G. (2004) *Agency and Answerability: Selected Essays*. Oxford: Oxford University Press.
- Wilde, O. (2003) *The Portrait of Dorian Grey*. London: Penguin.
- Williams, B. (1973) *Problems of the Self*. Cambridge: Cambridge University Press.
- (1993) ‘Nietzsche’s Minimalist Moral Psychology’ in *European Journal of Philosophy* 1 (1), 1-14.
- Wolf, S. R. (1993) *Freedom within Reason*. Oxford: Oxford University Press.
- Zagzebski, L. (2014) ‘Emotional Self-Trust’ in S. Roeser and C. Todd (eds) *Emotion and Value*, 169-182. Oxford: Oxford University Press.