

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

8-2022

Cultivating the human narrative: on Nietzsche, science fiction, and the aesthetics of life.

Derek Carpenter
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>



Part of the [Aesthetics Commons](#), [Continental Philosophy Commons](#), [Ethics and Political Philosophy Commons](#), and the [Film and Media Studies Commons](#)

Recommended Citation

Carpenter, Derek, "Cultivating the human narrative: on Nietzsche, science fiction, and the aesthetics of life." (2022). *Electronic Theses and Dissertations*. Paper 3959.

Retrieved from <https://ir.library.louisville.edu/etd/3959>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

CULTIVATING THE HUMAN NARRATIVE: ON NIETZSCHE, SCIENCE FICTION,
AND THE AESTHETICS OF LIFE

By

Derek Carpenter
B.A., Indiana University Southeast, 2014
M.A., University of Louisville, 2016

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in Humanities

Department of Comparative Humanities
University of Louisville
Louisville, Kentucky

August 2022

CULTIVATING THE HUMAN NARRATIVE: ON NIETZSCHE, SCIENCE FICTION,
AND THE AESTHETICS OF LIFE

By

Derek Carpenter
B.A., Indiana University Southeast, 2014
M.A., University of Louisville, 2016

A Dissertation Approved on

August 8, 2022

by the following Dissertation Committee:

Dr. John Gibson (Chair), Department of Philosophy

Dr. Ann Hall, Department of Comparative Humanities

Dr. Asaf Angermann, Department of Comparative Humanities

Dr. Vincent Joshua Adams (Outside Reader), Department of English

DEDICATION

This dissertation is dedicated to the memory of Wallace Earl Carpenter,
a true exemplar,
and to the memories of Winford Kenneth Cardin and Jerry Cardin,
whose absences have left this world a much less musical place.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my dissertation committee: to Joshua Adams for helping to bring important perspectives into this discussion on such short notice, to Asaf Angermann, whose enthusiasm for this project and whose depth of knowledge about virtue ethics has been inspiring; to Ann Hall, whose feedback has been crucial at every step, and who has taught me how to be a much better teacher; and to John Gibson, whose insights on Nietzschean ethics first motivated me to pursue this topic, and who has never hesitated to offer his invaluable guidance, support, and time. I would also like to thank Simona Bertacco and Michael Hagan, who believed in my ability to complete this project whenever my confidence wavered, and Elaine Wise, Pam Beattie, Robert Luginbill, and Leigh Viner, who were all instrumental in my progress to this stage. I would like to thank my wonderful and supportive family: Betty, David, Cheryl, Becky, Tim, Lindsey, Tammy, and Ken. Finally, I would like to thank my beloved Vicki, a brilliant autodidactic philosopher in her own right, who has been by my side from the beginning of this journey.

ABSTRACT

CULTIVATING THE HUMAN NARRATIVE: ON NIETZSCHE, SCIENCE FICTION, AND THE AESTHETICS OF LIFE

Derek Carpenter

August 4, 2022

In this dissertation, I will argue that there is a specific model of morality which can account for the intuition about moral features that we take to be intrinsic to the human type. This model of morality is that of Virtue Ethics, but not a conventional kind of Virtue Ethics. Instead, I will argue that the moral philosophy of Friedrich Nietzsche presents a model of cultivating narrative types that we take to be grounded in the kinds of beings we are, without resorting to any problematic metaphysical systems. This will require that I defend a novel interpretation of Nietzsche's ethics – a debate which is ongoing. In the first chapter of the dissertation, I will argue that Nietzsche's valuations can be reconciled with his attacks on conventional ethical systems, and that his value claims can best be understood with reference to his aesthetic views. In the second chapter, I will explore the specific details of Nietzsche's broad attacks on ethical systems, and then defend his subsequent value claims as analogous to a unique kind of virtue theory. In articulating how this kind of virtue theory functions, I will also propose that this is how we arrive at our moral conception of the human narrative. In the final chapter of this dissertation, I will draw from specific examples of science fiction narratives to

elaborate two specific views of the genre-specific trope of the “artificial human being.” In exploration of this trope, I find that our intuitions of the moral human are affirmed in our narrative fascination with ontological types, and also that there are under-explored implications of how these narratives are developed.

TABLE OF CONTENTS

INTRODUCTION.....	1
CHAPTER 1: AESTHETIC DIMENSIONS OF NIETZSCHEAN ETHICS.....	9
CHAPTER 2: IN DEFENSE OF NIETZSCHEAN VIRTUE ETHICS.....	52
CHAPTER 3: SCIENCE FICTION FILM AND THE HUMAN NARRATIVE.....	98
CONCLUSION: “A CONCENTRATED BUT SHIFTING VIEW OF HUMAN LIFE”.....	145
REFERENCES.....	152
APPENDIX.....	157
CURRICULUM VITA.....	158

INTRODUCTION

In the early part of the last decade, a series of short videos were released by the Boston Dynamics Laboratory, documenting advances in robotics technology. In these videos, engineers introduced various quadrupedal and bipedal robots, and then subjected them to what resembles repeated “abuse” – knocking them down and physically impeding their paths, or introducing other various ways to confuse and interfere with the completion of their assigned tasks. The intent of these videos was to demonstrate the resiliency of the machines in recovering from physical upsets, as well as their powers of computational reasoning in response to unforeseen resistances. Aside from the anticipated wonder at how machines could now operate at such sophisticated levels of autonomy in problem-solving, another reaction to these videos was surprisingly common – sympathy. Many viewers would react with pity toward machines which it is reasonable to assume could experience no suffering as a result of this perceived mistreatment. If these machines are not beings capable of suffering, then why do these videos so commonly provoke this kind of reaction?

In a 2020 study, Christoph Bartneck and Merel Keijsers explored a similar scenario. Test subjects were shown videos depicting different kinds of violence inflicted upon robots, and then a second set of videos depicting the same kinds of violence inflicted upon humans. Interestingly, when comparing the two sets of videos, test subjects

generally believed both scenarios to be equivalently wrong, whether the victim was human or machine.¹ Perhaps reactions such as these have something to do with the similarity of robotic machines to beings that actually do possess the capacities for reason and autonomy – beings that may actually suffer pain and frustration. However, there is little in the way of anthropomorphic resemblance in the Boston Dynamics robots that should provoke such a response. In many cases, their form is vaguely recognizable as a chassis of limbs and torso, but that is as close as it gets. Nothing implies that they experience anything like frustration, pain, or confusion. Still, the behavior that we do see, the struggling through purposive obstruction and hindrance, seeing their bodies upended, struck, and shoved down, may affect us, if even briefly, as something fundamentally human-like in expression – the perceived similarity of this activity to human struggling and suffering.

Consider another example. In a video created by researchers Fritz Heier and Marianne Simmel, several animated geometric shapes are shown interacting.² A larger triangle “follows” two smaller shapes – another triangle and a circle – around the screen. The shapes occasionally collide and appear to react to one another in various ways. At one point, one of the shapes appears to “hide” from the large triangle. The experiment reveals that when subjects are asked to describe what they see in the video, they usually construct a narrative which *anthropomorphizes* the shapes in some way.³ The larger shape is often interpreted to be the aggressor, terrorizing the smaller shapes, who panic

¹ See Bartneck and Keijsers (2020). It must be noted that when these subjects were instead shown videos of the “victims” retaliating against attack, they were more willing to view the humans as justified in fighting back against robot attackers than the other way around.

² Goldman (2013). Accessed at: <https://blogs.scientificamerican.com/thoughtful-animal/animating-anthropomorphism-giving-minds-to-geometric-shapes-video/>

³ Heier, Simmel (1944)

and flee, eventually escaping from the large shape. The Heier/Simmel experiment suggests that to anthropomorphize the behavior of objects, and even to construct narratives which explain this behavior, is a conceptual framework by which we make sense of phenomena. This probably tells us more about our preconceptions of anthropomorphic behavior than about the phenomena itself, and specifically, what we take to constitute right and wrong behavior of beings toward other beings. However, if we are affected in this way by such minimal degrees of resemblance to human activity, then we can easily see how we could be further affected if the degrees of similarity were much closer, as in instances of more elaborate and intentional mimicry of human action, such as those depicted in the narratives of science fiction.

When we examine these kinds of situations in narrative constructions – the robot or cyborg, the *artificial* human beings of science fiction – we find that in addition to the sense of futuristic wonder that they lend to such stories, they embody kinds of details that we value about our status as human beings, and facts about how we actually define that state through these values. We derive a sense of how our fundamental idea of humanity is differentiated from the hypothetical *not*-human, and the effects of this variability of difference. In this differentiation, we see how we *perform* the concept of the human, and in this performance, the features of the category of humanity which are crucial to our narrative of self are also revealed.

One of the most significant dimensions of difference between humans and the human-like machines of science fiction is how we understand human beings in relation to categories of morality. If a human-like machine is so different from us that it does not share any of the most fundamental capacities we associate with agency and moral

reasoning, then we may draw two conclusions: we do not hold this machine to any *moral* responsibility, and we do not owe any moral responsibility toward it. The reason that this claim seems strange is that it is so intuitive. There are no machines that we hold to a standard of moral responsibility, nor are there any that we are morally obligated toward, except for those obligations related to property rights, various terms of legal usage, and the like, but these exceptions are actually only kinds of obligations toward other human beings. Though I believe it is possible that this claim may age poorly, at the time I am writing this, no machine is currently entitled to what we could rightly refer to as *human* rights.

There is a moral sense of the human – an idea of how beings *should* act, based on the kind of beings that they are. Human beings hold a standard of moral responsibility that mere machines do not. If we do not hold non-human beings to a moral standard because they are incapable of moral reasoning, then our sense of morality is contingent to a state of being that we currently understand as uniquely human. We may get angry at machines, but we always understand that a machine is a kind of thing that cannot operate at a moral capacity. Similarly, we may expect animals to understand different levels of preferred behaviors, and we may even employ morally-loaded terminology that reflects this when they fail to instantiate the preferred behaviors, such as “bad dog,” or “ill-tempered cat,” but we do not actually hold animals to be moral agents, nor do we label them as transgressors or exemplars. To be human is then to inhabit a unique state in which morality is a fundamental reality for us.

But what about a hypothetical scenario (a thought experiment) that introduces machines that are *almost* human; would these beings warrant a greater degree of moral

responsibility than *mere* machines? Are there varying degrees of a concept such as human moral responsibility? How do we feel about immoral action toward a being that is obviously not human, but that we know is capable of experiencing the negative effects of mistreatment in the same way that we do (unlike, presumably, the Boston Dynamics robots)? How would we feel toward the fully human beings that would inflict this mistreatment, in full awareness of the effects on these hypothetical humanlike machines?

The differences I have discussed reveal consequences of our designation of a *moral* type of humanity. We think of human beings as occupying a unique state in which we can engage in questions about moral responsibility, and so there is something in the narrative of the human type that automatically entails some kind of moral dimension. Do we also believe the *right* kind of moral behavior rather than the *wrong* to be associated with humanity in the authentic sense? It is common for us to consider humanity in correspondence with specific moral features, such as an automatic inclination toward good moral behavior. For instance, we consider cruelty to be an expression of the *inhumane* dispositions of certain types of individuals. When we appeal to the goodness of individuals, we often say that we make an appeal to their *humanity*.

In the chapters that follow, I will argue that there is a specific model of morality which can account for our intuitions about moral features that we take to be intrinsic to the human type. We have a general sense of what constitutes good action, and what kinds of moral outlooks we take to be a *properly human* level of exemplary moral conduct, but we have problems in grounding these views in conventional moral systems. The reason for these problems is that our moral systems often push us into accepting contradictory directives or committing ourselves to unsound moral ideals. I will argue that the viable

alternative to flawed or inconsistent foundational systems lies in the moral philosophy of Friedrich Nietzsche.

In the first chapter of this dissertation, I will discuss Nietzsche's strategies of valuation, and argue that in his attacks on problematic moral systems, he offers us an alternative morality, reflected in the kinds of exemplar narratives that he develops over the span of his writings. Importantly, Nietzsche's moral views are best understood with reference to his aesthetic views. In other words, the kinds of behaviors that we should *normalize* are those which embody beautiful ways of living, on Nietzsche's terms. Identifying which specific behaviors and characteristics Nietzsche endorses can be difficult given the discursive, and even at times contradictory ways that he articulates his visions of exemplar types. However, Nietzsche's moral philosophy, unlike the conventional systems which he attacks, does not outline specific answers to the complete spectrum of moral challenges, but instead guides us to the point that we are able to supply our own answers to these challenges, free from influences which would constrain us to patterns of action which are stagnant, mediocre, or even grotesque. Generally, this means ways of living that demonstrate creativity, authentic commitment to self-expression, and some level of ability to recognize the value in *unconventional* perspectives.

In the second chapter, I will explore the specific implications of Nietzsche's broad attacks on ethical systems, and then defend his subsequent value claims as a more stable form of virtue theory. In articulating how a *Nietzschean* conception of virtue functions, I will also propose that this is how we arrive at our *moral* conception of the human narrative. Whereas conventional moral systems prescribe actions based upon presumptions of autonomy, intention, and the ability to anticipate the similarity of

outcomes with any kind of prescriptive relevance, Nietzsche observes that we do not have reliable access to these criteria.⁴ Instead, Nietzsche presses us to cultivate habits that will help to free us from these systems, including those dispositions toward independent thinking, and the ability to recognize different forms of the human narrative as valuable on these terms. When we conceive of those moral characteristics that we take to be essentially human, we have been conditioned to accept a specific set of virtues as the *right* type of human. By contrast, a Nietzschean virtue set would entail the cultivation of habits by which we create our own virtues, and our own narrative conception of the human exemplar. This Nietzschean conception of the human virtue set is no less a narrative construction than the conventional model, but it represents a more sustainable narrative, an aesthetic ideal.

In the final chapter *of this dissertation*, I will draw from specific examples of science fiction narratives to elaborate two specific views of the genre-specific trope of the “artificial human being” in order to establish the argument of a *moral model* of humanity outlined in the previous chapters. In exploration of this genre trope, I find that our intuitions of the moral human are affirmed in our narrative fascination with ontological types, and also that there are under-explored implications of how these narratives are developed. In one version of the human narrative, the *non-human* are cast as flawed, or even monstrous by comparison to their authentically human counterparts. In the second version of the narrative, conventional distinctions between *human* and *non-human* are challenged as arbitrary, revealing that what we value in the exemplar types is not essentialized in the form of biological human beings, but in a range of moral behavior

⁴ BGE §32

that we quantify as uniquely human, opening the door for a range of human types that may not be, strictly speaking, human. While the first narrative type presents the excessively optimistic view that *good* exists in us simply by virtue of the type of beings that we are, the second narrative type acknowledges that those characteristics we take to be uniquely human are actually the result of an ongoing process of self-cultivation – of refining the human narrative. In showing us a “robot’s progress” from machine to authentic humanity, many science fiction narratives provide a metaphorical view of our own existential condition, of constantly becoming, and overcoming the narrative of self.

CHAPTER 1

ON THE AESTHETIC DIMENSIONS OF NIETZSCHEAN ETHICS

1.1 Introduction to Nietzschean Ethics

The most prevalent theme in Nietzsche's writings is the general upending of knowledge structures which ground conventional systems and values. Because of this, it is easy to conclude that there is no room in his philosophy for discourse frequently affiliated with normative thinking and theory. Some critics argue that to interpret Nietzsche as a moral theorist risks misrepresenting his overall intent, which is to free us from ethical structures, rather than substitute new ones in their place.⁵ Further complicating the notion of a coherent moral strategy is Nietzsche's rejection of conventional models of agency, rendering the designation of moral responsibility problematic or impossible. It is challenging to conceptualize how Nietzschean normative claims or theories might be grounded without contradicting these criticisms, though one observation remains clear: alongside Nietzsche's attacks on stagnant, outmoded, and ill-founded value systems he presents their antitheses — both elaborately articulated judgments and implicit appraisals of what kinds of actions and beliefs constitute a life of worth and merit. In other words, Nietzsche's critique of conventional ethics creates an alternative to those systems.

⁵ Berry (2015), Foot (2002).

As with conventional ethical systems, in Nietzsche's writings, claims about actualizing a particular kind of life become prescriptive strategies. Throughout the discussion that follows, this project will demonstrate that Nietzsche articulates a unified ethical vision, and that uniquely situated as a consequence of his critiques, it is resilient to the threat of self-contradiction under the rigor of his own wide-ranging skepticism. The best way to demonstrate that there is such an ethical strategy in Nietzsche's work is through an examination of the intersection between his ethical, metaphysical, and aesthetic views. Such an approach will offer insights into how we as human beings formulate contingent sets of values, specifically those embodied in a narrative construction of human agency.

1.2 Nietzsche's Ethical Skepticism: Truth, Part One

One response to apparent incongruities in Nietzsche's work is to reject the interpretation of possible moral commitments in preservation of his more obvious deconstructive arguments; in other words, Nietzsche is not seriously engaged in the development and defense of ethical claims and is only concerned with discrediting the suppositions on which such claims may be grounded. If we interpret Nietzsche this way, we may argue that what is commonly ascribed to be ethical claims in his work are generally misinterpreted as such or are merely metaphorical and nothing more.⁶ However, this strategy is unconvincing. Nietzsche's outwardly discursive commentaries on culture and society consistently imply a state of affairs in contrast to that which is

⁶ In other words, to hold such a view would be to reject the views of those who interpret Nietzsche as a positive "moralist," including Robert Solomon (1985), Christine Swanton (1998), and Christine Daigle (2015).

reflected in standing cultural norms. Nietzsche constructs elaborate thought experiments of what human life could, and in many cases *should* resemble. Ideal modes of living are revealed to us in distinct narratives and aphoristic characterizations of a kind of human agent that has achieved mastery over a preferred range of existential possibility. Any such narrative presumes a background state of values; Nietzsche clearly elevates one way of living over another. In this way, specific values that are embodied in the *exemplar* agents of Nietzsche's accounts become prescriptive.

Where do we situate Nietzsche's presumed value claims in relation to his attacks on the foundations of ethics? It would be puzzling to suggest that Nietzsche (or any committed skeptic) would so comprehensively attack the weak metaphysical grounding of value norms and then proceed to build new normative arguments using the same kinds of grounding. Generally, Nietzsche does not reject the possibility of morality, but the standing systems by which we normalize moral practices.⁷ Reconciling Nietzsche's axiology with his general criticisms is possible. It is simply the case that one must proceed with a clear distinction between what Nietzsche is specifically attacking and that which he is defending or advocating. To understand how Nietzsche promotes specific sets of moral values alongside his widespread rejection of value systems, we must understand that the targets of Nietzsche's critiques embody morality in a way that Nietzsche finds both harmful and ill-conceived. The critical lens through which we identify how we value these values, and by which we subject our values to deeper scrutiny becomes a foundation for Nietzsche's moral stance. By subjecting the intuitive foundations of forms of morality to such scrutiny, many of them are revealed to promote

⁷ See BGE §7, "Our Virtues"

norms which are in conflict with one another, or which otherwise fail to produce beneficial outcomes.

Brian Leiter categorizes Nietzsche's collective references to these unsound, disadvantageous, and contradictory types of morality as "morality in the pejorative sense," or "MPS".² In Nietzsche's writings, MPS is often held to possess any of the following irredeemable flaws: a problematic assumption of agency and action which grounds the evaluation of acts as either "good" or "bad," an unsupportable identification of similarity in all occurrences of kinds of actions evaluated under MPS, or a general restriction of exceptional human beings to a level of mediocrity.⁸ The justifications for Nietzsche's rejection of these specific presumptions are thorough and I will explore each of them further in the next section. For the present discussion, the most relevant failing of MPS is that it erroneously presumes a world which is fundamentally accessible to the faculties of human reason, and compatible with the aims of human flourishing.⁹

In the most general sense, though they presuppose a great deal, ethical systems that fall under the umbrella of MPS entail sets of statements which, on the surface at least, are not controversial. These statements are supported by prescriptive arguments about how human beings should act under various circumstances. Claims such as "one should do this" or "these outcomes are ultimately preferable to others" conventionally rely upon knowledge of the nature of human agents and their relationship to the world in

² Leiter (2015). From here forward, following Leiter's terminology, I will continue to employ the abbreviation "MPS" when broadly referring to ethical systems that Nietzsche implicitly or explicitly rejects.

⁸ For Nietzsche's critique of agency, see BGE §32 and GM §1.12; for Nietzsche's critique of the identity of similar actions, see GS §335 and BGE §215; for Nietzsche's critique of values which make us mediocre, see BGE §228 and GS §21.

⁹ BT §3; HH §517

which they exist. Both deontological and consequentialist ethics ultimately draw ethical conclusions based on a recognition of what we take to be patterns of causation. In the case of Kant's deontological ethics, some *will* to act *causes* a subsequent action. In the case of consequentialist systems of ethics such as those of Bentham and Mill, this basic acceptance that we can recognize patterns of cause and effect entails a belief that we can anticipate that certain actions will *cause* reasonably predictable outcomes. An ethical argument of this kind is therefore a metaphysical one — meaning that if we can understand at a fundamental level what kinds of beings we are, what kind of world we inhabit, and how this world works, then we should be able to logically infer the kinds of circumstances that are most favorable for us, by an accepted measure of value, and then be able to prescribe actions that follow from this reasoning. Nietzsche raises a variety of challenges to all of the above presumptions, not the least of which is the question of whether the accepted metaphysical foundations of moral reasoning are actually available to us at all.

At the beginning of *The Birth of Tragedy*, Nietzsche acknowledges a philosophical intuition which is difficult to escape: the view that there is a part of the world separate from that which we directly perceive, and that owing to limitations of perception and reason, this “hidden” level of reality remains necessarily inaccessible.¹⁰ This intuition is reflected in the skepticism of many classical traditions, but also motivates the dualities explored in modern systems such as those of Kant and Descartes. The problem skeptics confront is that we only have access to appearances, and things

¹⁰ BT §1, §15

appear inconsistently to us at least some of the time.³ Skeptics warn that we may become reliant upon a kind of dogmatic belief that the world will conform to expectations we form based upon patterns of appearances. This problem is compounded when an argument is justified in principles which rely upon something more fundamental and enduring than appearances – i.e., the underlying nature or “essence” – that elusive layer of reality we intuit to exist somewhere outside of the reach of our senses. In trusting our faculties, we can never eliminate the risk that either our perceptions or the reasoning which follows from them may be inaccurate.

The difficulty of truth correspondence in the wake of skepticism is significant for Nietzsche’s critiques of ethics. Because we have no reliable access to the presumed foundations of knowledge, metaphysical or otherwise, we have no means of verifying that our argumentative claims bear any relationship to the foundational truth that they intend. Do universal systems such as ethics actually presume to be grounded in complete and accurate knowledge (an impossibility), and in formulating these systems do we *fabricate* truth, filling in the gaps where our knowledge runs out? Do we fill in these gaps in our knowledge of the world in a way that serves underlying agenda or intentions, rather than the interest of truth itself? Or more importantly, do we impose an interpretation which disguises some of these agenda?

To accept the belief in a fundamental level of truth which is distinct from, and may not actually resemble, our current experiences of the world is to accept that on some level all of the so-called foundational principles on which our knowledge (ethical or

³ See Sextus Empiricus, *Outlines of Pyrrhonism* §1; *Zhuangzi* §1.

otherwise) is grounded could be wrong, or worse, knowledge at this “foundational” level could be impossible. However, to formulate principles intending them as universal necessities presumes that these principles are valid under the full range of possible particular conditions. Even accepting that our present range of knowledge does not encompass all available facts of these conditions, faith in the metaphysical grounding of principles entails the notion that eventually, theoretically, it could do so. The more that we believe that we learn of the world, the more we are tempted by this possibility.¹¹ At any rate, we cannot bypass the skeptical concern that this kind of inquiry is doomed – that it would eventually presume the accessibility of knowledge that is forever and necessarily beyond our reach.¹²

Even more problematically, further inquiry into the nature of the world only increases awareness of the epistemic limits of the conditions within which we exist. The more we try to resolve our ignorance, the more ignorant we realize ourselves to be, and the more incapable of expelling this ignorance. If we are unable to discharge this skepticism about the limits of our knowledge, then it complicates the effort to justify ethical arguments in any kind of objective metaphysical necessity. At the very least, any inquiry directed at the honest and rigorous pursuit of foundational knowledge ultimately only pushes us toward the inescapable awareness of its own limitations. As Nietzsche observes in *The Birth of Tragedy*, “[...]science, spurred on by its powerful delusion [of

¹¹ In *The Birth of Tragedy*, Nietzsche states: “Anyone who has experienced the intense pleasure of Socratic insight, and felt it spread out in ever-widening circles as it attempted to encompass the entire world of appearances, will forever feel that there can be no sharper goad to life than the desire to complete the conquest and weave the net impenetrably close” (§15).

¹² GS §335

the accessibility of knowledge], is hurrying unstoppably to its limits, where the optimism hidden in the essence of logic will founder and break up.”¹³

If we accept the proposition that the divide between appearance and reality will never be closed through the process of our inquiry, then this also entails the tacit acceptance of a duality: there is the world of human experience, and the world behind it that we either do not or cannot experience. This means that the mundane realm of human experience, flawed as it is, characterizes a type of existence which we hold to be fundamentally illusory, in contrast with the realm of truth — objective and undebatable, necessary but unattainable. The distinction between truth and illusion, and the long-standing skeptic discussions of the discernible relationships between this duality serve as background to Nietzsche’s metaphysical discussions in *The Birth of Tragedy*.

If we take the basis of knowledge to originate at a level of reality below appearances – i.e., the so-called “world-in-itself,” which is necessarily outside of our realm of experience – then on what basis could we successfully evaluate human conduct or establish moral directives? The problem is actually more complex than it may seem, as the conventional foundationalist model presumes a knowledge of self which is likewise elusive. Once again, in ordinary ethical discourse we operate under the customary assumption of what kinds of beings we are and under what sort of circumstances we live in order to enact principles grounded in what is *good* or *bad* for us. But Nietzsche challenges this view, observing, “[...]there is no such substratum; there is no “being” behind doing, effecting, becoming; ‘the doer’ is merely a fiction added to the deed – the

¹³ BT §15

deed is everything.”¹⁴ To further emphasize this, Nietzsche proclaims that counterintuitively, “Everyone is farthest from himself.”¹⁵ In this and other similar passages, Nietzsche appears to adopt a view that a human being is not identical to a ground of consciousness from which agency emerges and action then follows; a human being is instead a particular kind of convergence of underlying yet indeterminate phenomena which express themselves through actions.¹⁶ As Aaron Ridley argues, for Nietzsche, the intention to act and the action itself are not discrete phenomena, but are merely parts of a single “mutually constitutive” event.¹⁷ In effect, Nietzsche collapses agency and action into a single phenomenon, and this is what is actually accessible to the scrutiny of our moral determinations.

While Nietzsche’s model of agency resolves the skeptical concerns outlined above by avoiding problematic strategies of valuation which are generated by or validated through foundational determinations, many of Nietzsche’s narratives of human agency employ language which appears oddly inconsistent with the postulates of his own theories. Frequently, Nietzsche encourages the existential freedom to forge our own accounts of self: for example, in *The Gay Science*, Nietzsche urges, “*What does your conscience say?* – ‘You should become who you are.’”¹⁸ Later in the same text, he writes, “We, however, want to *become who we are* – human beings who are new, unique, incomparable, who give themselves laws, who create themselves!”¹⁹ If we reject the

¹⁴ GM §I.13

¹⁵ GS §335; Nietzsche subverts the Socratic axiom, “Know thyself.” Editor Brian Williams also notes that this is an inversion of the German expression, “Everyone is closest to themselves.”

¹⁶ GM §I.13

¹⁷ Ridley (2018, 17)

¹⁸ GS §270

¹⁹ GS §335

notion of a “doer behind the deed,” then we may question where the agency which lends coherence to this project of existential creativity originates. What is this voice of conscience, and how do we give weight to its judgments, if it is not something we can infer to actually exist?

Alexander Nehamas argues that Nietzsche views our tendency to *reify* the self and the world as a kind of abstraction – a desire to re-conceptualize the *process* of “becoming” as the *state* of being (in this case, the *ego* self), and then to project this outward, and into the world.²⁰ Is it possible to accept that our models of self and agency are in a sense a *fiction* and still commit to their validity, and find them valuable on these terms? As Nehamas claims, in pursuing these abstractions, Nietzsche is not contradicting himself by positing that which he explicitly rejects – a fixed and stable *essence* of self. Nietzsche is only attempting to explore the implications of an existential habit that we are already engaged with – the process of “aestheticizing” life. For Nietzsche, the world that we engage with is analogous to a kind of literary narrative, and we are characters enacting its plot through different forms of living.²¹ This does not imply that such narratives, while lacking any attainable objective correlate, are without meaning or value. We accept some works of literature as fiction and are still moved by fictional elements, and are able to consistently evaluate some narratives as superior to others.

We can see where conventional prescriptive strategies of ethics go awry in attempting to reify dynamic processes into a static entity. In designating a specific account of the way the world is as the fixed and authoritative fact of the matter, the

²⁰ Nehamas (1985, 171); TI §III.5

²¹ Nehamas (1985, 5; 62); BGE §22, §230

corollary presumptions that follow from this static account are already out of sync with the fluid world of human experience.

As Nehamas argues, Nietzsche understands that all accounts of the world are necessarily *perspectival* and so the parameters of perspectives expressed in these accounts in turn reveal what we value in them. In fact, for Nietzsche, a “perspective-less” objective view of the world is conceptually incoherent (GM III.12). Every time we provide an account of the world, we must necessarily do so from our particular perspective, which is necessarily limited to a specific range of experiences. In describing the world, we must exclude something of it – that which is outside of our range of knowledge. But also, we choose a range of particulars from among our experiences and emphasize what we deem to be valuable to express in the account. These collective exclusions and inclusions comprise an individual perspective on human life. In the same way that an artist must limit their focus to which elements they would see represented in a particular work, every individual draws upon different preferences and patterns in constructing their own account of the world and pulling these elements together in a meaningful way.²² Living a particular kind of life is therefore a product of choices, and these choices reflect the value sets of agents. This process warrants scrutiny, as in some cases, these exclusions can constitute a kind of “epistemic violence” – the *oppression* of those viewpoints that are rejected from the narrative.

On Nietzsche’s perspectivist view, value norms and judgments arise out of particular views of the self as a kind of coherent fiction – a conventional account of these perspectives embodied in a narrative of agency. The search for the ground of agency

²² Nehamas (1985, 50)

which is *expressing* these perspectives leads us back into a regress, but this does not mean that we can or should discount the validity of perspectives on that basis. Our notion of agency is in a sense a fiction, but only in the same sense that any perspective of the world is. Nietzsche's rejection of the ground of agency and then subsequent invocation of the concept of agency is essentially a narrative of one who creates narratives. Rather than accepting the default narrative we are pressured into by the circumstances in which we live, Nietzsche urges that we express a will to create our own narrative of selfhood, and embody the kind of agent we *should* wish to be.²³ This implies a prescription of a specific kind of will toward agency. Agency is something that by definition requires active engagement with one's own activity; if we internalize the habits and rules of others, then strictly speaking, we are not agents, we are automatons.

The emphasis on where the *agency to create agency* originates becomes less problematic, as long as we recognize agency in this sense as untethered to the kind of metaphysical objectivity that MPS presumes. In any account of agency and the world, some features are affirmed as the most relevant aspects of a particular subject, and always represented as such from some specific point of view. The world is an intersubjective field of these perspectival accounts of *being*. Our narratives of self and world comprise a coherent understanding of what we take to be forms that particular ways of living may take. Value norms arise as we judge these narratives against one another. As abstractions of what truth we may actually know, we may consider the content of existential narratives, and examine our responses to them, unencumbered by the entanglements of global skepticism and metaphysical necessity. In short, narratives, in their range of

²³ GS §335

representing ways of living, are by default, means of communicating values.²⁴ This is not to say that Nietzsche proceeds as if there are no other constraints on his process of value formation, or that all narratives hold equal value.

How exactly do we quantify the moral content of a specifically *Nietzschean* existential narrative? If we embrace the freedom to create our own meaning through ways of living, within what moral parameters should we exercise this freedom, according to Nietzsche? The most consistent way that Nietzsche answers this question is rather simple: we should live in such a way that embodies a beautiful kind of “narrative” of life, rather than one which is mediocre, grotesque, or otherwise repellent.²⁵ Ultimately, the way of life that Nietzsche consistently advocates in his writings, and by extension the apparent objective of his moral thought, is the actualization of a kind of aesthetic ideal. Nietzsche encourages us “[...]to be poets of our lives, starting with the smallest and most commonplace details.”²⁶ Specifically, for Nietzsche, this entails ways of living that demonstrate a kind of existential creativity, authenticity, and the agency to create meaning. Nietzsche’s critiques of morality (to be elaborated further in Chapter 2) are nothing more than a rebuke of those systems which fail to instantiate a beautiful narrative of human life, as well as the tendency of society in its various forms to reinforce those stagnant systems. Not everyone is capable of fully exercising this kind of radical existential freedom, but for those who are, it becomes a kind of responsibility.²⁷

²⁴ Kupfer (2006, 341); GS §78

²⁵ GS §290

²⁶ GS §299

²⁷ GM §2.2; GS §301. I will address this argument further in Chapter 2.

1.3 Nietzsche's Aesthetics of Existence: Truth, Part Two

To interpret Nietzsche's moral judgments through the lens of aesthetics requires a minor shift in conventional axiology. Nietzsche frequently blurs the distinctions between categories of value, and in his writings artistic modes of evaluation become generalizing principles which contribute to a variety of ethical, metaphysical, and epistemological arguments. Consequently, in developing his critiques, Nietzsche's aesthetic ideas shed some light on questions normally reserved for ethical discourse, and good moral behavior often exemplifies qualities of artistic merit. However, this view is not entirely controversial; in ethical discourse, judgments are often expressed in aesthetic terms; good moral conduct exemplifies a kind of beauty.²⁸ As Daniel Came observes, the quality of "virtue" is not merely a reflection of the capacity to perform the right kind of moral action, but also a quality which stands as a possible object of aesthetic appreciation.²⁹

For Nietzsche, in the wake of the collapse of MPS, the aesthetic mode of reflection becomes the primary (and possibly only) available vehicle of moral evaluation. Given the failure of conventional normative strategies, we are more reliably justified in judging actions in the same way that we would evaluate a work of fiction or of visual art.³⁰ In this way, we may derive a conceptual exemplar which is both ethical and aesthetic – foundational to our notion of an idealized and abstracted self. We create the notion of a human being whose conduct embodies qualities of aesthetic merit, or alternately, fails to do so in some significant way. We use this exemplar as a point of comparison to our own conduct, and the conduct of others. The process of aligning

²⁸ Came (2014, 128)

²⁹ Ibid. 130

³⁰ Ibid. 128

ourselves with a “worthy” aesthetic paradigm is a highly stable way of grounding our moral commitments, motivating our actions, and imbuing our lives with meaning.

The idea of “living beautifully” as a moral end is not the extent of Nietzsche’s moral innovations, nor is this idea exclusive to his philosophy. Beauty is a concept which, no different than the moral idea of the “good,” is contingent upon specific evaluative frameworks. For Plato, a life lived in accord with principles most accurately reflecting the abstract metaphysical realm of the *forms* instantiates both moral and aesthetic ideals - albeit imperfect manifestations of the absolute perfection which funds our knowledge and experience of all things.³¹ Similarly, Aristotle locates the highest moral achievement in a kind of cultivated harmony between the world and the faculties of human reason by which virtue (ἀρετή) is developed and maintained. The exemplar of the virtues stands as an object of aesthetic reflection.³² In each of these cases, the object of aesthetic value follows from the demonstration of a specific kind of ontological perfection.³³ Simply, for both Plato and Aristotle, some level of human perfection is possible, and different degrees of this perfection manifest beauty to variable extent. Both good action and beauty are values measured against the objective model.

This kind of classical *eudaimonism*, the ideological traces of which extend well into enlightenment traditions and beyond, strikes Nietzsche as naively optimistic about the possible harmony between human beings and the world we inhabit.³⁴ To understand

³¹ Plato’s “Theory of Forms” — see *Phaedo* 100c-d; *Cratylus* 439b-c; *Symposium* 211a-c; *Republic* 507b

³² Aristotle, *Nicomachean Ethics*, (Hereafter referred to as NE) §X.6-8

³³ Kain (2009, 153). Aristotle’s *Megalopsychoi* may bear little substantive similarity to Nietzsche’s *Übermenschen*, but each presumes an account of a particular kind of being and their relationship to the world, and what the exemplary form of this being may look like on this model.

³⁴ Beiser (2006, 31)

why, it is helpful to restate a general summary of the *eudaemonist* thesis in terms of two related arguments: 1) The world is rational, and therefore, by extension, human beings are rational parts of it. Further, because of this harmonious relationship between human beings and the world, we are well-suited to the endeavor of discovering a way to fulfill our flourishing, by virtue of natural rational capacities to do so, and the availability and possibility of such flourishing within this world. Through a rigorous kind of intersubjective discourse (philosophy), we can more precisely determine principles and actions best in accord with these pursuits. Contingent to this argument, we may have some kinds of responsibilities (moral duties) toward others and in accommodation of their own pursuits of flourishing based upon these same facts. 2) It is in the best interests of human beings, as rational beings seeking a life in harmony with a rational world, to pursue the kinds of rigorous discourse outlined above and to accept the ethical and metaphysical foundations they presume. In fact, the philosophical tradition initiated by Socrates holds our suffering to be a necessary symptom of our ignorance regarding the true nature of these foundational concepts.³⁵ If classical *eudaimonism* holds knowledge up as an aesthetic ideal, then it follows that ignorance and falsehood are conditions which constitute aesthetic flaws.

The first argument of the thesis stated above faces the skeptical challenge: the nature of the world which Socratic/Aristotelian *eudaimonism* presumes may not actually exist, or if it does, we can never reliably determine it as a necessary truth-condition from which to derive normative arguments. While the skeptical challenge is a problem for the thesis overall, Nietzsche is more consistently critical of the corollary argument, which

³⁵ BT §15

entails the deceptively convenient notion that the capacities of human reason are well-suited to reveal necessary truths of the world, and that through this inquiry we may discover the ideals of moral and aesthetic worth, and thereby achieve the highest possible level of human flourishing.

Because Plato's Socrates professes that our suffering is related to our fundamental state of ignorance, he prescribes that we should systematically relinquish all illusions in search of the truth which they obscure, even as limited as this foundation of truth happens to be.³⁶ On the surface, this seems like a worthwhile objective, but for Nietzsche, the problem is that the promise of Socrates represents a more pernicious kind of illusion. We bury disagreeable truths only by replacing them with a lie – the aforementioned assertion that the search for truth will facilitate an end to human suffering.³⁷ But any reliably truthful inquiry will eventually reveal that suffering is inescapable, and that to believe otherwise is a reckless delusion.³⁸ As Nietzsche warns, even the advances of scientific inquiry, driven by the same epistemic optimism of Socrates, will reach its limits, and give way to the “boundless” frontier of the unknown, leaving only the “*tragic knowledge*” of our circumstances in its wake.³⁹ We cannot therefore pursue flourishing in the Socratic sense because we cannot consistently avoid meaningless suffering once we have unraveled the truth of its necessity. Ultimately, then, the Socratic account of the universe is both aesthetically and epistemically deficient; there is nothing aesthetically meaningful to be derived in its fiction, nor is there any well-founded truth to be reached in its inquiry.

³⁶ See the “Allegory of the Cave,” (*Republic* 514a–520a)

³⁷ In BT §15, Nietzsche advises that the Socratic tradition labels error as “inherently evil.”

³⁸ Gemes, Sykes (2014, 95); BT §17

³⁹ BT §15

Nietzsche warns us that any rigorous inquiry into the nature of the world, to the extent that this is possible, only appears to support a fundamental disharmony between ourselves and the world.⁴⁰ Rather than being well-suited for the pursuit of flourishing, our rational capacities may only lead us so far as the realization that the world is actually a horrible place for us, a place of suffering. In other words, there is no ontological fit such as that which the Socratic tradition presumes, and so rather than mitigating human suffering, the truthful inquiry may actually make suffering worse by reaffirming it as an inescapable condition of existence. The project of *eudaimonism* is therefore self-defeating. Consequently, as Socratic truth-seekers we may be deluding ourselves about the possibility of a world that is compatible with the ends and aims of human necessity and flourishing, leading us to confront the so-called “wisdom of Silenus,” which Nietzsche famously cites in *The Birth of Tragedy*:

“Wretched, ephemeral race, children of chance and tribulation, why do you force me to tell you the very thing which it would be most profitable for you not to hear? The very best thing is utterly beyond your reach not to have been born, not to be, to be nothing. However, the second best thing for you is: to die soon.”⁴¹

At this stage, in terms of Nietzsche’s metaphysical commitments, the pursuit of truth is at least problematic. An honest account of our condition only appears to reinforce the argument of *antinatalism* - the idea that existing at all is a bad thing for us.⁴² Worse, according to Nietzsche, the effort actually *has* succeeded in *falsifying* placating ontological narratives, including art, philosophy, religion, and myth — means by which

⁴⁰ In *Human, All Too Human*, Nietzsche states, “There is no pre-established harmony between the furtherance of truth and the well-being of mankind” (§517).

⁴¹ BT §3

⁴² See Benatar (2006)

life was previously made bearable.⁴³ Not only have we destabilized the epistemic security of these institutional structures, but we have come to recognize that our faculties for understanding the world are inescapably fallible.⁴⁴ In a modern (post-Socratic) world we must confront the growing awareness that the search for truth has dislodged the primary means by which truth itself has previously been justified as a worthwhile pursuit. In reality, complete truth is ultimately beyond our reach, and comforting fictions can no longer survive within the narrow range of truth that constitutes our experience. To continue to follow the path down which philosophy has led us is to face the twofold danger of a groundless metaphysics (and therefore ethics), and a life of suffering amplified by the growing awareness that this life is hopelessly meaningless.

The concerns outlined above motivate the existential questions at the heart of Nietzsche's philosophy: without access to truth, how do we derive meaning? Without meaning, how do we live? Is it possible to successfully refute the nihilism of Silenus? Fortunately, the incompatibility of truth and flourishing does not negate the possibility of a worthwhile life. It is simply the case that flourishing in the classical sense is not a realistic goal, given the conditions of existence. Nietzsche rejects the Socratic/Aristotelian focus on the synchronicity between truth and flourishing but does not reject that which he holds to be more fundamental: the value of the pursuit of meaning. For Nietzsche, it is not the case that we are incapable of enduring the necessary suffering of existence, but rather that we are incapable of bearing this suffering meaninglessly. As he states in *On the Genealogy of Morals*, "[humanity's] problem was

⁴³ BT §17

⁴⁴ GS §107

not suffering itself, but that there was no answer to the crying question, ‘why do I suffer?’”⁴⁵ Historically, the philosophical pursuit of truth has not answered this question, but only revealed it as unanswerable. This is a problem for us, as the kind of beings that can bear significant levels of adversity and unhappiness as long as there is some justification, some reason for our suffering. But too frequently, the further we investigate into the nature of our circumstances, we can only discover that there is no such meaning to be found, and no reasonable response to this “crying question” of suffering.

In *The Birth of Tragedy* Nietzsche famously pronounces that, “only as an aesthetic phenomenon is existence and the world eternally justified.”⁴⁶ What this means is that in art the question of suffering finds an answer. Suffering is made meaningful in the effects that forms of art have on the way that we process suffering. In art, we may conceive a vision of the world in which unbearable truth is mediated through an “aestheticized,” and therefore more manageable perspective.⁴⁷ In this way, we may frame groundless suffering in such a way that it is reconfigured as comprehensible, meaningful.

A tragic life is painful for the one who lives it, but there is also beauty in the art of the tragic narrative, and myth prompts us to recognize ways that such a life can exemplify this kind of narrative beauty.⁴⁸ In the tragic, and in acceptance of the horrifying truth of things, there is still an affirmation of our condition taking place, and we are moved by it. We may recognize the painful elements of tragedy as trivial when compared to the sublimity of the overwhelming reality it expresses. Nietzsche argues that in

⁴⁵ GM §III.28

⁴⁶ BT § 5

⁴⁷ BT §3; §18

⁴⁸ BT §9

mythologizing existence, the pre-Socratic narratives contextualize human suffering as a small but necessary feature of *cosmic* narratives.⁴⁹ In this way, the infamous declaration of Silenus stands as an account which casts the protagonist Midas as a tragic figure worthy of the audience's empathy. For human beings, even to endure meaningfully in the face of otherwise meaninglessness becomes its own noble form of living. While the *hubris* of Midas in demanding access to this knowledge reveals a kind of character flaw, as the audience we do not hope for Midas to succumb to Silenus' warning, but instead to embody a further kind of nobility inherent in a specific narrative of *exemplary* humanity — when confronted with the horror of existence to keep on living anyway. For Nietzsche, myths such as this illustrate some of the advantages we have sacrificed in the progress through Socratism toward Modernity. For Nietzsche, Socratism does not represent a refinement of cultural progress, but rather a decline; in the advent of Socratic thought, the cultural sicknesses which give rise to stagnant narratives and modes of thought are allowed to take root in society and proliferate.⁵⁰ MPS is a product of this decline.

If the best possible consolation for suffering lies in myth, then it would follow that, as an ordinary matter of practice, we should make effort to interpret and represent the world fictionally. But to what degree should we attempt this, and how exactly? Even if we accept the premise that what is best for us is to *fabricate* our own account of our reality, this raises numerous questions. If fictionalizing our existential narratives in the manner of Attic myth is accepted as beneficial to our well-being, then what prevents us from embracing delusional falsehoods completely, and devising fabrications far-removed

⁴⁹ Gemes, Sykes (2014, 83); BT §23

⁵⁰ BT (Preface, i)

from the accepted veracity of the so-called “factual” level of experience? There are certainly those who have gravitated in this direction, fabricating spurious narratives, often in order to affirm their own values or motives. Nietzsche is no less critical of those who conceal self-interested mores and beliefs within foundational narratives in order to exert power over others than he is of those who dogmatically accept and absorb them without challenge.⁵¹

How do we balance what we need to know with what would be “most profitable” for us not to hear?⁵² Fictionalizing the way that we view our conditions of living appears to involve a strange kind of self-delusion. Is this self-delusion even possible? Can we act based upon information that we already suspect may be false? Can we ever accept a state of affairs in which our actions are not guided by the belief in foundational knowledge, but by the acceptance of a fiction? Even if we hold that on some level we occasionally make sense of our world by filling in gaps in our knowledge in order to construct an intelligible narrative of life, we still generally prefer this narrative to represent *truth* in some significant way. A minor exaggeration or error would likely not disrupt the entire narrative, but extensively contrived or easily falsifiable premises might lead to its collapse, and the consequent problems of enduring life in the wake of such a collapse.

One such unsound narrative is that which Nietzsche refers to as the “ascetic ideal.” By this kind of account, human suffering is made comprehensible by introducing a universal “guilt” in which all human beings participate – sin.⁵³ Under this ideal, the necessity of suffering is not irrational, for God exacts a measure of universal justice, and

⁵¹ I will expand upon this claim in Chapter 2.

⁵² BT §3, the “wisdom” of Silenus.

⁵³ GM §III.28

our suffering serves as a kind of atonement. However, in the wake of the collapse of religious systems corresponding to the rise of enlightenment rationalism, suffering has once again become meaningless, and ontological claims have become ungrounded. Nietzsche's figure of the "lantern-wielding madman" warns us that in the aftermath of this upheaval, we are left to wonder whether all such claims are rendered useless:

What were we doing when we unchained this earth from its sun? Where is it moving to now? Where are we moving to? Away from all suns? Are we not continually falling? And backwards, sideways, forwards, in all directions? Is there still an up and a down?⁵⁴

The failure of systems such as religion and Socratism implies a growing awareness that our values cannot be grounded in any consistent or necessary truths, but only arbitrary associations.

1.4 The Apollonian, the Dionysian, and the Bridge Between

It is difficult to ignore the metaphysical intuition addressed earlier: that our experience of the world is reducible to a duality of truth and illusion. Conventionally, we understand truth to be equivalent to that which represents the absolute underlying reality of things, and the way that things appear to us may or may not actually correspond to this reality in any meaningful way.⁵⁵ Again, our unease arises from the concern that truth is ultimately either inaccessible or dangerous to pursue, and that illusion is founded on that which we cannot trust as reliable. This leads us to speculate about the most manageable balance between the two. To miscalculate is to risk reliance upon untenable epistemic foundations, to become misled by meaningless delusions, or worst of all, to actually

⁵⁴ GS §125

⁵⁵ BT §1

uncover the terrible and unendurable reality which grounds our experiences of things. Nietzsche gives us many examples in which societal norms incorrectly prescribe this balance, but what kind of narrative model entails a successful synthesis of *truth* and *illusion*, while still representing an *authentic* narrative of human life on Nietzsche's terms?

For Nietzsche, truth and illusion find optimal balance in the specific form of the pre-Socratic mythical narratives of Attic tragedy. As illustrated in the example of Midas and Silenus, Attic myths frame the underlying truth of human existence as a representation of our relationship to a superseding cosmic order. According to Nietzsche, through myth we confront the ongoing dialectic tension between truth and illusion, and it consists in the impulse to embrace and affirm the most honest account of our relationship to an overwhelming world (the *Dionysian* impulse) competing with the impulse to temper that account in tolerable layers of fictionalization (the *Apollonian* impulse).⁵⁶ In the most successful balance, these two primal forces complement one another – the *Apollonian* impulses restraining the *Dionysian* from a completely self-destructive awareness of the human condition, while the *Dionysian* revitalizes the rigidly safe and ordered *Apollonian* narratives. If such a state of harmony were to be achieved, we would not suffer excessively in knowledge, nor would we stagnate in repulsive states of self-delusion. This is not to say that such a harmony would resolve in a static condition – on this account the tension between these forces is dynamic and ongoing, a constant intrusion of one upon the other.

⁵⁶ BT §1-3

Harmony is an aesthetic ideal, but it is not one that exists in the world *prior* to our appraisal of it, at least in the sense that classical ontology presumes. In response to Nietzsche's rejection of the possible harmony between human beings and the world; there is no such underlying harmony because there is no orchestrated harmony within the world itself. We are merely elements of an indifferent network of dynamic processes, and as such, we are subject to its random nature. We are susceptible to the potential harm of states of disorder, as subjects capable of experiencing suffering. It is comforting for us to *psychologize* the world of morally indifferent processes, or to otherwise render the meaningless meaningful by imposing a rational order that is not actually present. Mythical narratives acknowledge the discord of the world and render it as something relatable to human experience, framed in familiar ways.

As functions of their respective roles in mythical narratives, the gods bear specific characteristics. Apollo is closely associated with the "plastic" arts, such as sculpture and architecture, and as such represents the influence of precision and measure. In the *Apollonian*, the world is rational, and ordered into intelligible categories. Conversely, Dionysus is embodied in drunkenness, divine "madness," and euphoric religiosity, as the manifestation of the unpredictable, irrational processes at the heart of nature.⁵⁷ The essence of the *Dionysian* influence lies in those "dangerous" and disagreeable truths that we may wish to shield ourselves from in the *Apollonian*. We may be too often tempted by the unrealistically optimistic view that our world should disproportionately reflect

⁵⁷ BT §3

expressions of the safely ordered and intelligible *Apollonian* rather than the dangerous and disordered *Dionysian*.⁵⁸

The problem, Nietzsche suggests, is that in withdrawing from the *Dionysian*, we recede further into self-delusive narratives, but self-delusion is a limited consolation, as it is easily refuted through our experience. It is evident to us that human existence is not a safe or painless state of being, otherwise the insular security of previously collapsed existential narratives would never have developed to harbor the individual to begin with. In contrast to *Apollonian* narratives, Dionysus does not promise to shelter us with happiness or to free us from suffering in his ecstatic departures; he instead promises access to authentic existential truths. Central among these truths is the awareness that to exist is to regularly face one's own destruction, and to acknowledge this is to affirm the pull of the *Dionysian*.

1.5 The Aesthetics of Existence: Beyond Truth & Illusion

Nietzsche's philosophy, while related by common threads of a unified critique of conventional structures, often reflects a shifting and discordant view of the relationship between art and truth, and the impact of this dialectic tension on our conceptions of human life. Similarly, the aforementioned "apparent incongruities" in Nietzsche's moral thought can be interpreted as representing a progress through different stages of his confrontation with the project of truthful inquiry, and the possibility of this project under the conditions in which we exist.

⁵⁸ BT §8

The idea that we temper our experience of reality with some degree of fictionalization in order to make life more bearable implies access to a reality against which our accounts may be measured as fictional. If some account of the world is *more* or *less* truthful or fictional, then there must be a foundational level of experience which stands as the criterion of truth, and we must be able to know something substantial about it. However, we cannot commit to the metaphysical accuracy of any account of the world if our truthful inquiries have only proven that our faculties for processing and understanding the world truthfully are inescapably fallible. Beyond *The Birth of Tragedy*, a notable shift occurs in Nietzsche's metaphysical commitments which appears to acknowledge this problem. Instead of approaching *aestheticized* accounts of the world as interpretations which falsify or fictionalize an objective truth about the way the world actually is, Nietzsche begins to suggest that individual accounts of reality are importantly *constitutive* of that reality. This refinement in Nietzsche's view acknowledges that the world is comprised of a plurality of individual perspectives, and that none of them are more privileged in terms of their correspondence to an objective truth of things.⁵⁹ Once again, though, this does not imply that all narratives hold equal value for us, when measured against aesthetic criteria.

If all that we have access to are the stories that we tell ourselves – about ourselves, and about our world — then what does it mean to “suffer” from the truth, and to need to fabricate a new framework by which to experience this truth? There is no substratum of reality that gives rise to the various metaphorical interpretations we impose onto it, whether we designate them as accurate or erroneous. What we hold to be truth is

⁵⁹ Nehamas (1985, 3)

always a perspective — truth from a particular view.⁶⁰ Nietzsche’s shift in metaphysical commitments involves a move beyond questions such as “what is truth?” and “Is there truth at all?”, leading us to the question of which “truths” are worthy of accepting. However, a *perspectivist* view does not mean there is no value in the world; it means that the value that we derive is not a necessary feature of the world which is independent of the subjective experience of that world.

This leaves us with an apparent dilemma – how can a perspectival account tell us anything important about the world if it has no truthful connection to that world? The answer: in a purely objective sense, it cannot. We have already established that a metaphysically necessary “truthful” state of being either does not exist or is too far beyond the scope of our experience to inform any meaningful account of it. This is actually not as great of a problem as we may take it to be. An existential perspective is a means by which we render experiences coherent, typically in the form of a narrative of meaningful living. The world for us is just a practical and coherent integration of experiences that we may accept to be “caused” by that which is outside of the realm of our experiences. We seem to be the kinds of beings that are generally very good at generating successful frameworks by which to make sense of these experiences, and to anticipate relationships between them. When we worry about the objective truth of existential narratives, what we really worry about is that some of them are wrong and that we may commit ourselves to them anyway. But with no objective grounding by which to privilege one account over another, all accounts are equal in terms of epistemic justification.

⁶⁰ WTP §481

Accepting the simultaneous validity of a plurality of existential narratives presents challenges. How do we rate one perspective as superior to another? If no account of the world is most demonstrably accurate in representing it, then why would we advocate for it instead of others? The same holds true for moral valuation; how can we appraise the actions of one individual as preferable to those of another if we have no means of claiming one to be objectively morally grounded? If every way of living reflects a kind of narrative grounded in the values expressed by an agent, then we must worry about moral relativism, which is a short path to moral nihilism. Relativism is a challenge to the proposition that there is such a thing as value, that our norms are grounded. Clearly Nietzsche is no relativist, as a driving focus of his philosophy involves arguing the inferiority of various conventional beliefs and normative systems to his own views.

Once again, the aesthetic paradigm offers a possible response to this problem. For Nietzsche, we are reliably justified in evaluating our actions (collectively constitutive of an account of human life) in the same way that we would evaluate a work of art.⁶¹ This may be a strange proposition for us to accept because, among other concerns, it seems logical to assume that impressions of a “fictionalized” world would have significantly less normative impact than a truthful account of the world, whatever we take that to be. However, aesthetic reflection is actually an effective mode of action guidance because it entails specific aversions to properties which we hold to be repellent or grotesque.⁶² In terms of an aesthetic account of human life (a narrative), this aesthetic reflection extends to include the habits and behaviors of agents (characters). On this model, we effectively

⁶¹ Came (2014, 128)

⁶² It is important to note that aversion does not necessarily entail avoidance.

become the audience and critics of our own existential narratives.⁶³ We can evaluate the perspectives represented in the form the narrative takes and determine whether this kind of life embodies aesthetic value.

Even if we acknowledge the aesthetic model as an indirect grounding for prescriptive ethics, it may seem problematic that, hypothetically, nothing discourages us from outlandish exaggerations far removed from, if not objective truth, then that which is commonly accepted to be the truth of things. There are two possible responses to this concern that are consistent with Nietzsche's views. First, many examples of this kind of "exaggerated" narrative have features that render them flawed on a number of aesthetic criteria. These flaws may include behaviors which are either so out of sync with the circumstances of the narrative that they contribute to its overall incoherence, or behaviors which are motivated by concealed interests which are themselves aesthetically odious. Alternately, that which appears to us as a radical divergence from an anticipated narrative pattern could actually represent the kind of radical existential creativity which Nietzsche praises. The field of existential possibility is unforgivingly indeterminate, leaving us to speculate what choices we should make in establishing aesthetic-existential frameworks.

If we intend to understand what it specifically means to us to judge existence on aesthetic grounds, then we need to elaborate which specific ways of living hold aesthetic merit for us, and which ethical strategies are flawed on this basis. In conceiving of life as a coherent narrative, we assume a more or less complete picture of the world, making it accessible to our ethical and metaphysical speculation. In a sense, human beings are both the artists as well as the work they create. For Nietzsche, it is the artists who, like

⁶³ Came (2014, 139)

physicians diluting a remedy until it is more tolerable, know what “fictionalized” elements to combine in forming this picture, and in what measures.⁶⁴ We come to value artists as those who are “experts” at successfully identifying and representing something meaningful in the world.⁶⁵

In *The Birth of Tragedy*, Nietzsche equates the philosopher’s relationship to metaphysics with the artist’s relationship to the illusory state of dreams.⁶⁶ Whereas the philosopher posits narrative accounts of existence which intend truth, the artist is comfortable with accepting illusion. Both are *creating* an account of reality. What distinguishes Nietzsche’s perspectivism from his earlier views is that in adopting an existential perspective, we are not strategically deluding ourselves by masking truth with illusion; we are embracing a truth. In this way, the roles of philosopher and artist collapse into one. Philosophy leads us to become artists of existence, the *exemplars* of creating exemplars.

The puzzling part of Nietzsche’s analogy is that it returns us to the *Apollonian/Dionysian* dialectic of truth and illusion. But in acceptance of the awareness that any presumed truth can only ever be an expression of perspective, it is difficult to reconcile these views without softening the commitment to truth as metaphysical necessity. If there is no perspective or narrative that bears any necessary relationship to foundational truth, then what do the *Dionysian* and the *Apollonian* impulses really represent? I have discussed how the *Dionysian* impulse implies the existence of a truth over and above whatever narrative we adopt in order to make life meaningful and

⁶⁴ GS §299

⁶⁵ Janaway (2014, 55)

⁶⁶ BT §1

therefore sufferable. But this is itself a narrative perspective of existence — the idea that we endure the unendurable by effecting some agency to exert control over it and make it meaningful through our own powers of existential creativity, ideally without withdrawing into the supposed safety of excessive self-delusion.

Nietzsche argues that we exist immersed in an ongoing dialectic tension between impulses toward truth and impulses toward illusion, but the meaning of this conflict is contingent upon how we understand these forces to actually represent features of our existence. We may understand the dialectic as an existential narrative of authenticity — a perspective representing the conflict of an unrestrained truth of existence, suppressed by pressures toward a safe and structured mode of living. On this view, the tension as previously stated still stands: *Apollonian* illusion is excessively optimistic about the world and *Dionysian* truth is excessively pessimistic – at least from the perspective of individuation of the self, and neither on its own lends itself to the exemplary narrative of existence. Nietzsche’s narrative of authenticity portrays the *Dionysian* impulse as more reflective of reality than the *Apollonian* illusion by juxtaposing an account of life lived according to standing conventional frameworks with existence forged in clear defiance of them. The narrative prompts us to see the truth “between the cracks” in the standing narrative, to question what is actually being repressed by the comforting veil of illusion. Through the way that we navigate this balance, we have access to the possibility of living meaningfully.

Society itself is founded and maintained on a variety of narratives. On one such narrative, the principal function of “civilizing” pressures consists in the repression and elimination of some natural human tendencies with the general intent of establishing and

preserving an ordered (and presumably thereby safer) type of existence for all participants of a society. Because human beings are on some levels a danger to themselves and to one another, systems of rules and privations have developed to restrict some actions in the interest of our own preservation. By such an account, social organization is merely a pragmatic bargain for us; in civilization we are bound together in a communal defense against the elements of the natural world which would endanger our survival. In exchange for this kind of insurance against extinction, we relinquish the unbound freedom of some of the most destructive of our primitive impulses.⁶⁷ This narrative holds that such *rational* restrictions are imposed upon individual liberty in order that specific patterns of social organization may endure, and social harmony may thereby be preserved. A tension may also arise, however, when the actions of some individuals diverge, and threaten the social harmony of this structure from within. Society must therefore maintain not only a level of resistance to harmful external forces, but a code of laws which aim to prevent the disruption of its internal order as well.

Some existential narratives embody our concern that society's system of privations extends too far and fundamentally limits something authentic in our expression of the narrative of human nature. The worry is that effects of systematized repression only amplify the very dangers they seek to minimize.⁶⁸ In the *Bacchae* of Euripides, Dionysus stands as a threat to the rational realm of society: he is the part of the world that is still unfamiliar, the *unheimlich* – formless. *Dionysian* formlessness is itself an affront to one of the most fundamental presuppositions of the *Apollonian* narrative — the

⁶⁷ See Freud (2010)

⁶⁸ Nussbaum (*Bacchae* introduction xv-xvi)

principle of individuation. Nietzsche speaks of the “blissful ecstasy which arises at the collapse of the *principium individuationis*.”⁶⁹ The erosion of conventional differentiations between self and world, as well as between self and other selves, may be interpreted as destruction to those whose conception of existence is still grounded in the *Apollonian* narratives of society. But as the balance shifts away from the *Apollonian* perspective, the *Dionysian* subject revels in the advance of this collapse of identity. The weakening of barriers of individuation and differentiation is represented as a kind of healing, and acceptance of a more honest ontological account — one that acknowledges human nature as being at least partially at odds with the civilizing pressures of society’s imposed categories.

However, when the barriers of reason fall away and orderly *Apollonian* structures begin to reveal their inadequacy, it is not merely the case that human beings are unleashed from the constraints of society’s network of rules and privations, but that they are forced into the orderless unknown. Anything that is known relies upon a conventional distinction between that thing itself and that which is not that thing.⁷⁰ The state of the unknown is so to us only because it represents the region where what we know runs out, where that which defies categorization is now the prevailing reality. The form that the god Dionysus takes fluctuates, and so it can be said that his essence is the absence of any fixed essence.⁷¹ This itself is a subversion of society’s categories, which intend toward universally and metaphysically grounded principles, but cannot admit the particularity which stands outside of their prescriptive order. A completely rational model of the world

⁶⁹ BT §3

⁷⁰ Versenyi (1962, 88)

⁷¹ Ibid. 84

will always demonstrate its systematic inefficacy in the failure to account for exception. The *Dionysian* is the particularity at odds with the *Apollonian* universal, and in the *Dionysian* the inescapable flaw in the system is made more obvious to us.⁷²

In the *Bacchae*, it is through the arrival of Dionysus on Earth that the mandates of society are disrupted, and that the pull of the irrational subverts the order of existing rational structures. The *Dionysian* is revealed as *other* in juxtaposition to dominant normalized institutions and ideologies, which are likewise only revealed in their difference through this dialectic. The narrative of the *Bacchae* also suggests that through this tension, an underlying natural state of being which is primordial to societal institutions is restored in the re-emergence of the influence of Dionysus himself. When Pentheus, King of Thebes, attempts to have the newly emerged Dionysus bound and imprisoned, Dionysus proclaims, “I am sane. I will not be bound by the insane.”⁷³ Dionysus continues, “You don't even know who you are.” This scene implies a reversal of a more conventional view of wellness and “normality”; for the God Dionysus, to deny the “insane” element within us all is clearly the true state of aberration. Dionysus accuses Pentheus of ignorance of, or worse – of willful rejection of his own nature. In driving the occupants of Thebes insane, Dionysus has not corrupted them, but reconciled them with a more honest ontological narrative, encouraging characteristics that were already present and giving them free reign.⁷⁴ From this perspective, Dionysus is not a destroyer, but a restorer, and the civilizing pressures of society are what are afflicting us.

⁷² I will expand upon Nietzsche's critique of the “universal” in the next chapter.

⁷³ *Bacchae* 33

⁷⁴ Versenyi (1962, 86)

The *Bacchae* narrative expresses a perspective of unease directed at the structures of civilization. The disruption of the presumed stability of social order may only be amplified through a failure to acknowledge some element of disorder inherent in individuals, and through regulations that bind groups of individuals to a flawed narrative of social order. Dionysus reminds us that we cannot hide from the truth behind an *Apollonian* facade and go on living as if the occluded “truth” is not still there waiting for us. Pentheus, in attempting to interfere with the influence of Dionysus, represents that repressive force in society, and his ultimate fate reflects the danger of overzealous restraint of primordial *being* in conventional individuation.⁷⁵ As a meta-narrative framing our relationship to truth through fictional narrative form, *Bacchae* illustrates an inescapable reality: that some fundamental truths cannot be concealed indefinitely. The existential perspectives which embody aesthetic value are likely to be ones which frame existence in a way that yields a beautiful narrative, but not ones which embellish our experiences with naive optimism. Whatever narrative that we embrace will require some rhetorical mechanism to account for the essential discord which appears prevalent in both individuals and social structures.

Society relies upon an array of standing universal norms for its function, and pressures individuals to condition themselves to operate successfully within these norms. To deviate from these norms often entails punishment or stigmatization. We may accept that some normalized restrictions are beneficial, but social systems also often restrict the vital innovation which is necessary to prevent the stagnation of the greater societal

⁷⁵ In the *Bacchae*, Pentheus’ mother Agave is driven mad, along with the rest of the women of Thebes; she is deceived into dismembering Pentheus in her delirium, believing him to be a wild Lion.

narrative when new cultural conditions outgrow old practices. Innovation is usually punctuated at an individual level, and social structures are slower to evolve. Some individuals display unique capacities to step outside of culturally conditioned modes of experiencing and interpreting the world, and in the process, expand our collective norms beyond sets of outmoded or regressive conventional practices. Through their actions these exemplar individuals in turn originate exemplar narratives — more refined perspectives on human life, and on what it in fact means to be a human being under the circumstances in which we live.

Returning to the existential-aesthetic model, we find that there is something aesthetically favorable, or *beautiful*, in what we consider to be good moral character, and something aesthetically flawed in what we consider to be bad moral character.⁷⁶ Neither of these are valenced according to standing universal principles of good and bad, but are contingent to the context of a moral narrative, which is in turn contingent upon a complex of perspectives which comprise more broad cultural narratives. Some properties of character which would drive an individual to find value and meaning outside of standing cultural norms of valuation may hold positive aesthetic merit analogous to those of artistic innovation. Alternately, stubborn commitments to outdated and purposeless norms are likely to prompt negative aesthetic valuation, as if in response to art which lacks originality.⁷⁷ Nietzsche's exemplars are those individuals who embody creative ways of living primarily by advancing new ways of exploring the meaning behind the concepts that ground our existence. In this way, cultural innovations are perpetuated,

⁷⁶ Came (2014, 130)

⁷⁷ Came (2014, 137)

emulated, and the propagation of previous ethical norms eventually weakens. It is evident that Nietzsche consistently elevates the aesthetic paradigm of the nonconformist – the philosopher for the “day after tomorrow.”⁷⁸ It is also understandable why this paradigm usually involves a disruption of the conventional norms and values of today. A variety of existential narratives have grounded our ethical traditions, and for Nietzsche, many of these express questionable degrees of aesthetic worth. Recognizing in some general sense which narratives Nietzsche dismisses, what can we say specifically about those that he advocates? While Nietzsche does give examples of what the exemplary narrative might look like, his writings are generally more concerned with guiding us toward the radical freedom to create our own narratives.⁷⁹ Nietzsche does not establish himself as the author of the new narratives, but as a kind of mentor for their future authors. If we subject the values that are installed in different existential narratives to scrutiny, can we be sure that they are not values which perpetuate mediocrity?⁸⁰ Do they conceal the agenda of *ressentiment*?⁸¹

I have commented that the defense of a “Nietzschean” ethics is a contradictory kind of endeavor however we approach it, and so some points still warrant further elaboration. We should recognize what ethical systems are generally supposed to do and determine whether Nietzsche’s aesthetic framework can actually carry this weight, and if it can, we should clarify how.

⁷⁸ GS §335; Z §1.3-9

⁷⁹ GS §335

⁸⁰ GM §1.12

⁸¹ BGE §260

1.6 The Nietzschean Exemplar Paradigm

In the most general sense that we can identify, what are our ethical obligations to one another? In what principles are these obligations grounded? How can we know these principles? How can we be sure that there is any necessary connection between our ethical obligations and the principles grounding them? Ethical systems usually resolve such questions by taking some form of metaphysical inquiry for granted — often a presumption of essences and agency by which to derive principles which ground normative mandates. The problem of skepticism destabilizes this foundationalist view; systems that appeal to metaphysical grounding can never resolve their necessity. Either we are wrong about the existence of such grounding facts, or we could be right about them but never able to verify it. A better way to state the problem is this: if there is no access to such a ground as foundationalist claims presume, then why do these claims matter? Why should we care about their mandate? Ultimately, conventional ethical systems cannot answer any of these questions with the conclusive objectivity that they purport.

Recall the earlier critique of those conventional ethical systems which must problematically resort to *essentializing* in order to justify their claims — this is the view that what should be determined as moral behavior is contingent upon what is good for human beings, which requires the further assumption of what a human being actually is. However, with no access to the metaphysical grounding of “essences,” our notion of the human being is limited to a conceptual model — a value-laden perspective; in other words, the human being is itself a narrative construction. This is true for any strategy we

use to fill in the indeterminate question of being, whether it is one which claims objective metaphysical grounding or not.

In ethical discourse, we consider the moral content of our circumstances in abstraction and then posit prescribed actions based upon these considerations. Because of this, every ethical scenario we devise is a kind of narrative. As this discussion has addressed, our existential narratives hold variable degrees of value for us, most reliably measured in terms of aesthetic criteria. If our ethical norms are grounded in contingent aesthetic values rather than a prescription of action based universal rules or duties, this bypasses the problems of metaphysical grounding, but what normative functions take the place of this grounding? A valuable narrative may entertain us, or may move us to an even deeper kind of aesthetic experience, but why and how could it guide our actions? How do we approach these kinds of claims and judgments aesthetically? What actions are beautiful? By what means do we communicate or popularize which kinds of actions constitute an existence which is beautiful?

Accepting Nietzsche's view that every valuation of action necessarily emerges from a particular perspective of human life, and the consequent claim that our very notion of humanity is grounded in kinds of aesthetic narratives, we may still question whether Nietzsche's accounts entail any sufficient normative component to qualify as ethical arguments. What kind of ethical claims and judgments are aesthetic, and what kind of aesthetic models have the normative powers of an ethical argument? Nietzsche writes of the consequences of choosing various modes of living, but it is unclear whether he issues any specific calls to act upon the values expressed in his writings. In blurring conventional categorical distinctions, is Nietzsche applying his aesthetic moral judgments

prescriptively or only describing the underlying creative forces which impel us to express various ways of living?⁸²

Though Nietzsche's claims are contingent upon his own perspectives, in them there is still a tacit presumption of how specific kinds of exemplar human beings tend to act under specific conditions. Such exemplary actions may embody kinds of aesthetic merit, inspiring us to emulate them, compelling us to act in specific ways, to embody ways of living that are "beautiful" or otherwise engaging. Even if this is the case, and in light of Nietzsche's departure from objectively grounded systems, how do we determine that specific aesthetic values in existential narratives are meant to ground our conception of *moral* good? A feasible response to this question is that we are inspired to act by properties of character – in this case, the character of exemplars. The kinds of character traits that inspire us to act are generally those which consistently result in ways of living that we believe to embody valued narratives of existence. This description easily corresponds to a classical definition of the concept of virtue.⁸³

Recall that Nietzsche's "expressivist" model eliminates the customary differentiation between agency and action: "[...]the 'doer' is merely a fiction added to the deed — the deed is everything."⁸⁴ The collapse of this distinction effectively bypasses problematic appeals to many of the indeterminate or inaccessible elements which hinder conventional moral determinations. The overlap of Nietzsche's expressivism with his perspectivism leads us to a framework of moral valuation which is primarily aesthetic,

⁸² GS §290; Came (2014, 134)

⁸³ In *The Nicomachean Ethics*, Aristotle elaborates the concept of virtue at length by describing actions which are demonstrably successful and explaining behaviors which have been habituated by those who not only display a capacity for such action, but also a disposition toward it.

⁸⁴ GM §I.13

and functionally consistent with the classical model of Virtue Ethics.⁸⁵ On the classical model, we understand virtue as a behavior which exemplifies “good” action. The concept of virtue is “self-reinforcing”; we come to identify good character with good action, and vice-versa.

Because a narrative of ontological justification – i.e., what kinds of beings we are and under what circumstances we live – lends meaning to our standing conventions of how we should act, we are drawn to narrative forms which attempt to supply details concerning what it means to be human. When we quantify this supposed meaning inherent in the concept itself, our intuitions reveal that what is most fundamental to the concept is not in its essence or form; instead, when we refer to “humanity” we have in mind a specific range of properties, conceptually analogous to a virtue set. This virtue set of humanity entails properties of character that can be either cultivated or neglected, and so an exemplar type of human being is an ongoing project of moral development, not an automatic or accidental status derived from one’s biological condition. Consequently, what we designate as a human being is not actually a stable or fixed essence – it is a model of striving toward a standard of being which is reflected only in the kinds of stories we tell about human beings, both the exemplar types and those that fail to meet the exemplary criteria. Our “narrative experiments” concerning the nature of the human being are an imposition of meaningful existence in the face of unforgiving meaninglessness. In revealing narrative beauty through the ethical exemplar, the concept

⁸⁵ To further defend Nietzsche as a kind of virtue theorist, some additional distinctions must be made in order to account for differences between Nietzsche’s worldview and those of classical *eudaemonist* ethics, as well as the content of his moral claims.

of humanity is imbued with a transitory but necessary kind of worth. In a sense, we are always refining this narrative and its presumptive guiding principles.

Nietzsche's thought experiments provide some direction for understanding this process, and specific formulations of the exemplar type of the human narrative. In the exemplar life we find meaningful ways of living, but not necessarily a path to individual flourishing. Nietzsche's "free spirits" embrace the perspectival nature of their existential condition and embody the radical freedom to express their role authentically and meaningfully in the "tragic narrative" as it unfolds.

CHAPTER 2

IN DEFENSE OF NIETZSCHEAN VIRTUE ETHICS

2.1 Introduction

In blurring the conventional distinctions between branches of axiology, is Nietzsche developing prescriptive moral claims, or merely ruminating on the aesthetic dimensions of different particular modes of living? The difficulties of reading Nietzsche as a moralist are numerous, given his self-proclaimed status as an “extra-moralist.”⁸⁶ Though he regularly argues the advantages and disadvantages of various ways of living, he is often vague about whether these judgments entail specific calls to action, or if, as a result of his views of agency, we are even capable of acting upon them in any meaningful way. In the last chapter, I explored and defended the theory that a deeper understanding of Nietzsche’s aesthetic views resolves many of the challenges associated with reconciling his advocacy of specific values with his expansive rejection of our intuitive justification of value. As other scholars have noted, in his articulation of the relationship between art and life, Nietzsche proposes an alternative grounding for a unified strategy of moral valuation, but this interpretation of Nietzsche’s ethics is only partial.⁸⁷ In order to fully explain how Nietzsche’s aesthetic models function in a normative capacity, we must

⁸⁶ BGE §32

⁸⁷ Proponents of the aesthetic interpretation of a Nietzschean ethics include Daniel Came, Alexander Nehamas, Christopher Janaway, Ken Gemes, and Chris Sykes.

explore further implications of Nietzsche's conception of agency. In bringing these two approaches together, we can proceed with a clearer understanding of how we form value and create existential narratives.

Accepting Nietzsche's assertion that "only as an *aesthetic phenomenon* is existence and the world eternally justified," more must be said about the way that this aesthetic justification functions.⁸⁸ While Nietzsche maintains that in order to thrive in the collapse of obsolete systems we will require a special kind of existential skill-set by which to derive new ethical conventions, and that the value of these conventions should be based in aesthetic criteria, two rather obvious questions follow from this: what constitutes a criterion of aesthetic merit, and who should determine such criteria? A beautiful existential narrative may serve as a model of an exemplary kind of existence, but which standards of beauty is it that are *supposed to* inspire us to follow their example? It is perfectly plausible that those who engage in behaviors that we hold to embody morally repugnant kinds of lifestyles and worldviews believe them to be beautiful by their own aesthetic standards. Given the diversity of aesthetic predilections, we must still worry about being left with no useful means of distinguishing who is right about which ways of living are beautiful and which are not. The aim of this argument is to demonstrate how Nietzsche can legitimately promote an aesthetic form of ethics without risking a backslide into relativism.

Nietzsche often makes it clear that the best and most proper ways of living align with his own values, but what is it about Nietzsche's exemplar types that situates them as the arbiters of value? While many of Nietzsche's critiques involve dismantling our views

⁸⁸ BT §5

of objective universal standards of “good,” there still must be something in our actions which holds as necessarily valuable, if there is to be any inter-subjective value in our values. In Nietzsche’s writings, those who occupy the crucial (albeit self-prescribed) position as adjudicators of the exemplar narratives appear to possess some commonality of character which allows them to more easily diverge from the paths ingrained in us as humans living in the wake of modernity. In the various forms that Nietzsche’s accounts of exemplars take, these individuals are normally represented as those who overcome culturally conditioned adherence to the outmoded norms which comprise failing systems, such as the ones that fall under the umbrella of Leiter’s *morality in the pejorative sense* (MPS).⁸⁹

Our values arise from particular perspectives of human life, but without some common standard, we may still question whether Nietzsche’s arguments entail any sufficient normative component to qualify as ethical arguments. If there are no values which are common to all, then value is relative, and our moral prescriptions are trivial — every ethics an individual reflection on one’s own preferences and commitments. On the other hand, in Nietzsche’s criticism of ethical systems, we find that much of what we generally take to be universal and necessary is merely contingent — accidents of our circumstances rather than essential features of existence. But is there anything that does not fall into this category? How do we engage in a critique of the flawed conception of the universal while allowing for a margin in which to construct ethical principles with actual binding efficacy? Following Nietzsche, we may find that this margin is scant, but

⁸⁹ As explained in Chapter 1, Brian Leiter categorizes Nietzsche’s collective references to unsound, disadvantageous, and contradictory types of morality as “morality in the pejorative sense,” or “MPS”.

not absent. It turns out that the only universal values we can reliably accept are related to facts about agency — what it means to perform actions at all. In human agency, there is a bedrock of universal value in the form of Nietzsche’s conception of the *will to power*. Facts about what it means to *will* power provide clues as to where our values originate, and lend weight to our ethical claims, to the extent that this is possible. In Nietzsche’s exemplar narratives, the standard of aesthetic value, and the intrinsic value of willing power overlap.

In order to expand upon a viable model of agency as it relates to Nietzsche’s ethics, I should first clarify some basic frameworks. More broadly construed — what function do conventional ethical systems serve, and to what degree do Nietzsche’s strategies align with this? Further, what are the precise failures of MPS, and through what mechanisms are Nietzsche’s valuations more firmly grounded in valid arguments about agency and aesthetic merit? Any response to these questions should sufficiently address my claims about the function and formulation of Nietzsche’s exemplar narratives and their relationship to the “human” virtue set, the consequent claim that our very notion of what kinds of beings we are is grounded in kinds of aesthetic narratives, as well as the way these narratives extenuate the aforementioned epistemic and metaphysical challenges associated with the presumed foundations of morality.

In *Beyond Good & Evil*, Nietzsche surmises that the history of ethics can be divided into three broad phases: *pre-moral*, *moral*, and *extra-moral*. In the earliest phase, which he terms the “pre-moral” phase, value judgments were established purely with regard to the outcome of our actions, and with no concern for determinations regarding the actions themselves, or what sorts of causal mechanisms motivated them (such as

intentions).⁹⁰ According to Nietzsche, in the progress from the first phase (the “pre-moral”) to the second phase (the “moral” phase), we have abandoned consequences in favor of intentions as the absolute criterion of value. While the *pre-moral* characterizes the largest phase of our moral development — the entire pre-history of our existence on this planet — we are currently in the moral phase. Nietzsche argues that we have reached the point in our progression as a species that we must make the transition to the third phase — the “extra-moral.”⁹¹ All of our current ethical traditions are various manifestations of the second phase — the moral. What will this proposed third phase be like? For Nietzsche, it will be everything that the *moral* has failed to be, in terms of reliable means to ground value. Importantly, having moved past our prioritization of outcomes, in the extra-moral we will now abandon our focus on *intention*, and prioritize that which is *unintentional*.⁹² How do we prioritize that which is unintentional? Even more perplexing — can we *intend* to adopt a means to prioritize the unintentional? To understand how Nietzsche resolves these challenges requires a deeper understanding of the failings of the moral phase which have prompted this transition to the extra-moral, and of how an axiology based on the unintentional will save us from these failings.

2.2 Nietzsche’s Confrontation with the Moral Phase

In most branches of ethics, basic types of ethical arguments that we can make involve claims intended to establish how well kinds of actions conform to some criteria of right or wrong. This statement may not seem controversial, but it suggests that an

⁹⁰ BGE §32. Note that this “pre-moral” phase should not be identified with the later rise of Utilitarian ethics in the enlightenment period, which presumed a prescriptive calculus that could not function without reference to both actions and their causes.

⁹¹ Ibid.

⁹² Ibid.

ethical claim presumes access to further knowledge of criteria that ground conceptions of right and wrong, as well as how this process of justification actually works. The criteria by which we determine what is *right* can take many forms, as can the manner through which we reach these criteria.⁹³ In other words, a moral judgment does not occur in a vacuum, but carries with it the baggage of its whole tradition. Thomas Brobjer argues that, “One of the most common dichotomies made in respect to moral judgments is that they must be based on either the consequences of an act or the intentions of the acting person.”⁹⁴ Nietzsche’s identification of the moral phase with the criterion of *intention* cannot strictly speaking be exclusively associated with the second disjunct of Brobjer’s dichotomy; even when we employ moral theory in a way that prioritizes outcomes over intention, we still conceive of organizing our actions with the *intent* of producing favorable *outcomes*. As Brobjer notes, Nietzsche rejects both criteria, as each strategy intends a level of necessity which is untenable.⁹⁵

As the familiar paradigm of intention-based ethics, an analysis of Nietzsche’s critiques of Kant’s *deontological* system broadly demonstrates the overall fallacies of this kind of ethical approach. Under Kant’s ethics, determinations about our moral failures and successes are not contingent upon the success of our intended actions. Instead, Kant holds an action to be good if it is undertaken with a *good will* – that is, with the intent to

⁹³ For example, see Pojman, Vaughn, (2007, 1); Bourke (1968, §1). The basic principles of conventional morality address such questions as the nature of the criteria that serve as justifications for the types of ethical claims that we can make: is “the good” a transcendent value? Does it derive from religious conceptions of an afterlife and punishment or reward? Is it contingent on an abstract realm on an abstract conceptual realm (as in Plato), or is it related to the physical or otherwise concrete circumstances of one’s existence?

⁹⁴ Brobjer (2003, 64)

⁹⁵ Ibid.

perform the proper action in accordance with prior universal, rational principles.⁹⁶ For Kant, these rational principles are the sole criteria by which our actions should be measured. Kant's fundamental guiding principle is the *categorical imperative*, the moral proposition which advises us to "Act only on that maxim whereby thou canst at the same time will that it should become a universal law."⁹⁷ Subsequent to this claim, Kant also holds that our moral obligation to the categorical imperative is distinct from our volitional obligation to *hypothetical imperatives*, which impel us to act toward the fulfillment of some contingent needs or desires; the categorical imperative constitutes a moral obligation toward a good which is *good* in and for itself only.⁹⁸ For Kant, everything we need to know about our moral obligations can be determined as necessity, owing to a shared human *nature* in the form of universal structures of consciousness — the capacity to reason the same way in response to a full range of moral challenges.⁹⁹ Kant's moral philosophy faces a variety of obstacles. Here I will focus on some of the most relevant formal challenges to Kant's ethics that Nietzsche employs. While Kant is not the only foundationalist which falls under the threat of Nietzsche's criticism, his influence on post-enlightenment morality cannot be understated, nor can his culpability as an architect of MPS.

The moral necessity of the categorical imperative is based upon the problematic assumption that our capacity of reason is robust enough to interpret universal truths at a fundamental level, and to do so in such a way as to exhaust any contingency that his ethics may encounter. Faith in Kant's categorical imperative entails the belief that we

⁹⁶ Kant, *The Metaphysics of Morals* (Hereafter abbreviated as MM) 4.393-4

⁹⁷ Ibid. 4.421

⁹⁸ Ibid. 4.414

⁹⁹ Ibid. 4.405

should be able to recognize the good in establishing a particular action as a moral obligation in every possible situation. Any exception would be enough to frustrate our intention to *will* that some action should be prescribed as a universal law. However, Nietzsche convincingly argues that we should be mistrustful of the proposition that human reason and the world are so conveniently compatible.¹⁰⁰ We face the likelihood that our inquiries cannot fully reveal the most fundamental truths of existence, nor even determine whether such truths exist. There may always be aspects of human experience that *originate* with conditions which are inaccessible to our understanding. Ethics can only ever anticipate a partial perspective of this experience. Even if we were to accept that moral good consists in acting from a good will, we can never derive maxims by which to prescribe actions to be universal duties if our knowledge of the world, and even ourselves, is necessarily incomplete or flawed. In other words, Kant's view entails faith in sets of claims that are too strong for the foundations that support them.

Owing to necessary limitations in individual outlooks, any moral account that we may provide is necessarily "provincial", grounded in an incomplete objective picture of the totality of particular subjective human experiences. This premise is damaging for the case of an action-guiding categorical imperative which appeals to rational universal principles that we believe to be deducible *a priori*, and designate as binding to every human being by virtue of a shared rational capacity to reason along these same lines.

Unfortunately, social structures and their foundational ideologies are too often the basis of such rigid maxims that resist the acknowledgment of exception and cannot adjust

¹⁰⁰ BT §3; HH §517

when confronted by it. For Kant, for an action to be ethically prescribed means that it must always be so, regardless of contingent circumstances which may affect the outcome for good or bad. The admission of exceptional cases would invalidate the necessity of the entire structure; ethical concepts are not universal if they inform human beings what to do in *almost* all cases. Yet it is the very possibility of contingent exceptions that highlight the efficacy of Nietzsche's criticism of universal systems.

The world as a metaphysical ground – the “thing-in-itself” reality – is simply not an accessible part of our experience, and so we have no way of dismissing the possibility of exceptions, or of inconsistency in individual particular experiences. Even within the range of those experiences, exceptions to what is conventionally absorbed into the conception of the *universal* are likely substantially more prevalent than Kant's absolutist ethics allows.

Nietzsche observes that even if we could claim that particular practices were demonstrably ethically justified in some contexts, this does not establish their necessity in all similar circumstances, nor does it establish that we could ever anticipate every contingent circumstance that might affect our ethical judgments under those specific conditions. This is because the similarity between these separate events is likely only superficial.¹⁰¹ Not only are distinct events unique and subject to their own evaluative conditions, but they may also reflect a multiplicity of overlapping or conflicting moral commitments. In other words, we cannot proceed with certainty from knowing the *rightness* of one action to knowing the rightness of another action which we deem to be

¹⁰¹ GS §335

similar. We may apply our experience of precedence in making moral determinations, but in doing so we must accept that the similarity of precedent circumstances is only conventional and lacks the necessity that would be required to ground a universal rule such as the categorical imperative.

For Nietzsche then, a more effective and accurate moral perspective would somehow accommodate the dynamic and often *discordant* nature of our experiences – even those that bear some surface level of similarity – in order to be more useful to us. In *Beyond Good and Evil*, Nietzsche presents an analogy of the variability of our moral commitments as planets whose orbits are determined at different positions by different stars in their proximity. Under different circumstances, the same action may be evaluated differently.¹⁰² Aristotle reaches a similar conclusion when he equates the flexible norms associated with justice to the malleable rule employed by builders in measuring stone.¹⁰³ In much the same way that a more rigid measuring device would result in an inaccurate measure of the stone's contours, any measure of moral value which does not bend to accommodate the imprecision of human life will fail to provide any useful normative instruction. It is therefore not a flaw in ethical principles if they must evolve to adapt to the particularity of new circumstances. For example, under many moral systems, to tell a lie with the *intent* of preserving the life of another rather obviously carries a different kind of moral obligation than to lie with the intent of advancing one's own material interests. Moreover, in many cases, the value of actions is contingent on the way that intention and outcomes are linked together. Whereas Kant wants to bypass the

¹⁰² BGE §215

¹⁰³ NE 1137b32

contingency of outcomes (which can never be anticipated with necessity), the prioritization of intention may lead us to cases in which this prescription of universality is arbitrary or ineffective.

A system of universal rules can never adequately map the dynamic nature of human existence and must rely upon the unsupportable presumption of regularity and predictability in order to hold any normative efficacy. We should therefore approach static ethical systems with skepticism, insofar as they do not reflect the world that they presume in a realistic way. But if there can be no identity between two events and each is in fact distinct in its particularity, then the “top-down” application of universal action-guiding principles is only grounded in an approximation of the way things are, and not a necessity. Even more problematically, the quantification of a normalized conception of the universal often falls under the power of social institutions that exploit this power to express the aims of their own agenda.

However, there is another problem. To discontinue reliance on universals leads to a further problem. If ethical norms are not grounded by any supervening principles which are necessary and universal, then they are susceptible to the criticism that their prescriptive power is weak or trivial — far from useful as the universal maxims that Kant intends. In response to this criticism, it is important to note that there is a range of normalized principles, approximating the function of universals, which still serve us in a practical way, even when their grounding falls short of the universality they presume. We in fact do rely upon the similarity of distinct events to render an approximation of the world which is valuable to our everyday process of moral reasoning. Judith Butler argues that the overall usefulness of our ethical systems is complicated by the fact that our

formulation of them remains necessarily incomplete, but this does not render them necessarily useless.¹⁰⁴ We can only employ ethical theories which institute our most complete available approximations of what would ground our guiding principles, were we to ever discover them. As such, our ideas of right and wrong, and our judgments of value relating to these principles, are susceptible to ongoing revision, as they grow closer to actualizing better and more effective formulations of moral concepts. Ethical norms may evolve through refinements in application, and through adjustment to particular circumstances. Accordingly, when we evaluate our ethical commitments to unfamiliar circumstances, we may make adjustments and subsequently decide whether we are improving their normative efficacy in doing so. Through a repetition of this process, over time we may come closer to actualizing a useful set of particular norms, while discarding others along the way. The fallibility of a necessarily limited “perspectival” existence means that testing particular moral beliefs against those of others is a way of maximizing their reliability, to the extent that we are capable of doing so.

Normalized behaviors help us to develop and refine existential narratives. Ideally, retaining norms means that we have tested them against discourse which has reinforced their reliability over narratives which were not as resilient under such scrutiny. Kwasi Wiredu argues that shared rational discourse helps to make our beliefs and arguments more resilient to skepticism.¹⁰⁵ If there is a shared medium by which beliefs are communicated and evaluated against one another, then there is an improved chance that we can avoid committing ourselves to practices which presume an unreliable grounding.

¹⁰⁴ Butler (1996, 46)

¹⁰⁵ Wiredu (1980, 210)

Acknowledging that all knowledge claims are necessarily limited by perspective, the strategy of taking a larger range of perspectives into account amounts to casting a wider epistemic net. However, Wiredu's argument is itself limited by his presupposition of a "normal" cognitive range of human rationality which serves as the objective medium for shared discourse. Even ignoring the problem that this argument is still vulnerable to the same challenges of the skepticism he confronts, Wiredu's claim does not allow for significant deviations from a common biological mode of cognition.¹⁰⁶ Similar to the earlier critique of Kant, Wiredu's biological essentialism amounts to quantifying a limited range of types and reifying them as if they are objectively grounded. Nevertheless, Wiredu is correct that we cannot discount the importance of insights gained through discourse, contingent or otherwise, as without them, we could not evaluate the continued usefulness of our opinions.

As skeptics, we can still function with the awareness that there is no necessary relationship between a "world-as-such" and the structures through which we interpret it. It is not a problem for us that any such relationship is only conventional. The world that we quantify in language and in reasoning is only ever an approximation aggregated through a vastly elaborate system of metaphors which we may only ever communicate to one another approximately.¹⁰⁷ This is necessary because we must adduce a similarity between unquantifiable multitudes of disparate particular experiences in order to understand them, and more importantly, to make them meaningful. We create meaning by imposing these relationships between phenomena. If one particular experience

¹⁰⁶ Wiredu (1980, 178)

¹⁰⁷ According to Madan Sarup, "All language contains metaphor. Understanding metaphor is as much a creative act as creating metaphor." (1993, 48)

acquires a meaning for us, another experience is potentially intelligible by invoking the similarity it has to that original experience. However, recall Nietzsche's criticism of Kant's imposition of identity where only similarity may exist.¹⁰⁸ The dilemma Western epistemology cannot rid itself of is that it sets a criterion for knowledge which can never be achieved, and yet knowledge claims within the tradition are no less ubiquitous for the effort. But we do not need *a priori*, necessarily grounded truths to function; we need to create meaning. We can do this by approaching customary relationships between actions and outcomes and determining some to be valuable.

As Aristotle claims, "The rule for what is indefinite is itself indefinite."¹⁰⁹

Accordingly, when we apply our customary habits and interpretations to new sets of circumstances and find them wanting, the improper response is to cling to them resolutely out of deference to familiar "necessary" systems. Kantian ethics is unequivocally opposed to this kind of revision in the face of contravening evidence. The categorical imperative resists revision because it is instituted as if it already infallibly encompasses a complete range of moral criteria. In practice, if inconsistency arises, that which does not fit systemic categorization is often simply ignored or discarded. Societal structures assert a monopoly over value categories by claiming a necessary affiliation with the universal, and every particular that falls outside of categorization is presumed to be flawed, incomplete, *other*.¹¹⁰ But when we structure our categories in such a way that they intend the universal, they become removed from experience – they become something abstract, transcendent.

¹⁰⁸ GS §335

¹⁰⁹ NE 1137b30

¹¹⁰ See Foucault, *Truth and Power*

Herein lies the problem with ethical systems. We may always be confronted with particularity which falls drastically outside of that to which accepted norms are capable of prescribing a response, and to which previous experience can provide any reliable guide. For us, the world consists of an unpredictable array of conditions which may challenge the necessity of our moral claims. The problem therefore lies in determining whether we can conceive of a binding ethics which admits the possibility of revision, and honestly acknowledges the nuances and irregularities of human experience. Ultimately, as Nietzsche observes, no system has proven itself conclusively more reliable than any other, and so we are better off cutting ourselves free of ethical maxims as *monolithic* standards of value.¹¹¹

As I have mentioned, drawing a distinction between *morality in the pejorative sense* (MPS), and *morality itself* is important. Nietzsche's attack on the "transcendental" metaphysics which grounds universal structures does not extend to preclude the possibility of normative moral arguments, though they obviously must take different forms than the conventional models that he attacks. The solution to this disparity lies in distinguishing *principles* from *structures*. Robert Solomon outlines terminology which helps to clarify the distinction. Solomon identifies "morality" with those universal systems of norms that Nietzsche explicitly rejects, and "morals" as "generally agreeable or acceptable traits that characterize a good person."¹¹² For Solomon, a further distinction is made between these terms and "ethics," which refers to the broad field of discourse involving questions of morality and morals, and the relationships between. Solomon

¹¹¹ BGE §202

¹¹² Solomon (1985, 253); Lester Hunt makes a similar distinction, in discussing the difference between "big M" Morality and "small m" morality (1993, 23).

observes that for Nietzsche, the primary ethical question has to do with what the relationship between morals and morality actually is.¹¹³ The question is whether in the *extra-moral* phase which Nietzsche urges us toward, we will still have something that resembles morals. Solomon concludes that the *post-moral* Nietzschean morals, however we articulate them in terms of their content, are best understood as analogous to those of Aristotle, and antithetical to those of Kant.¹¹⁴ Aristotle offers us a blueprint for forging value norms in the face of the dynamic, contingent, and often *inconsistent* world – virtues. Virtue as a moral concept provides a way to understand Nietzsche’s transcendence of the moral phase, and his elevation of a moral criterion that is *unintentional*.

2.3 Ethics and Agency – A Move Toward the “Unintentional”

Even as Nietzsche reflects on the presumed advantage of the historical shift toward intention-based prescriptive ethics in his figurative genealogy, he already persuades us that a further move is warranted: a progression beyond intention as the criterion of moral value. Acknowledging that we routinely judge actions based upon what we presume to be the causal conditions of agency, Nietzsche claims that our identification of *intention* as the cause of our actions is inadequate – that the true cause lies in what is *unintentional* in the act.¹¹⁵ This claim is less paradoxical than it may seem. Any prescription that we act with a “good intent” requires meaningful knowledge about agents and their intentions. Under conventional models of the *moral* phase, an agent *will*s

¹¹³ Solomon (1985, 254)

¹¹⁴ Solomon (1985, 254)

¹¹⁵ BGE §32

an action, and we subject the content of this “willing” to moral scrutiny. However, as Nietzsche consistently argues, we have no reliable access to the internal structures of consciousness which moralists claim to ground our knowledge, and which would allow us to presuppose ethical maxims *a priori*.¹¹⁶ Investigation into these causal mechanisms leads us to a regress of hidden motivations, impulses, and other factors which are likely obscured to us.¹¹⁷ In other words, either there appears to be something in the causal sequence underlying agency which is prior to that which we commonly accept as its ground, or we are mistaken in believing that we can identify the reasons we act in any such way at all. Either way, that which we take to be “conscious reasoning” in the constitution of agency is not a stable ground for moral criteria.

It is ambitious to question something as fundamental as the supposition of an “I” that oversees all of our actions — the identity of self with agent, and agency with the process of willing, under the direction of conscious intention. But as Nietzsche points out, this impression we have which casts the agent as cause and the action as its effect requires an oversimplification of phenomena in order to fulfill this purpose.¹¹⁸ In *The Gay Science*, Nietzsche elaborates how we typically conceive of the role of the will in a presumably intentional action, such as striking an object:

[...]the feeling of *will* seems sufficient to him not only for the assumption of cause and effect but also for the faith that he *understands* their relationship. He knows nothing of the mechanism of what happened and of the hundredfold fine work that needs to be done to bring about the strike, or of the incapacity of the will in itself to do even the tiniest part of the work.

¹¹⁶ e.g., Kant’s Categorical Imperative

¹¹⁷ GM §1.13

¹¹⁸ GS §127

As we see in this passage, according to Nietzsche, the simple view of will and agency is reinforced by some feeling we come to associate with the act of striking, and these affective phenomena belie the fact that the processes involved may actually be more complex than we take them to be. Commonly, we conceive of the will in a very linear way; we *will* actions, and they subsequently occur, and so we reasonably conclude that we have directed the whole process. But conventional understandings of agency also allow that this may work the opposite way as well — sometimes we speak of the will as something which overtakes us and impels us to act. If we can be affected by the will in this way, then we are not identical to the will, nor are we even necessarily its architect and overseer. So then what exactly is the *will* and what is our relationship to it? If that which we take to be the will has some sway over our actions independently of our conscious intent, then somehow, we are beings that both will action and are willed to act. Brian Leiter observes that Nietzsche’s arguments about agency address the fundamental question of why we tend to identify ourselves with the phenomena which control the will to act, rather than those which are controlled by it.¹¹⁹ In his discussion on the nature of the will in *Beyond Good and Evil*, Nietzsche explains:

“Freedom of Will”— that is the expression for the complex state of delight of the person exercising volition, who commands and at the same time identifies himself with the executor of the order — who, as such, enjoys also the triumph over obstacles, but thinks within himself that it was really his own will that overcame them.¹²⁰

We generally have little problem accepting that some of our bodily activity is the result of unconscious drives or impulses. What is more difficult for us to accept is that

¹¹⁹ Leiter (2015(a), 121)

¹²⁰ BGE §19

these same kinds of internal or unconscious phenomena may be causally intertwined with what we take to be our *agential* activities — our deliberate intentions to carry out actions in fulfillment of our goals. The reason, according to Nietzsche, is that it feels better to conceive of ourselves as being in control of these phenomena than being controlled by them. Because we are the kinds of beings Nietzsche claims that we are, this “emotion of supremacy” that we come to associate with agency is a state that we enjoy — we thrive on the feeling of power.¹²¹ The preference for this state leads us to exaggerate the fiction of the *I* — the *doer* that precedes the *deed*. However, our preference for being in control is not sufficient to make it so. Our distortion of this phenomenon leads to the strange “doubling” effect discussed above — a branching of a single event into two: doer and deed.

Unconscious impulses and *drives* to act may be no less integral to the process of agency than our conscious faculties of reason. In fact, many more of these impulses and drives may be constitutive of agency than we apprehend or acknowledge, accounting for the obscure “hundredfold fine work” to which Nietzsche refers. Even those conscious thoughts which we take to motivate our intentional actions are in turn motivated by a multitude of unconscious mechanisms.¹²² Patterns of actions, behaviors, character traits — in fact, principle truths about who we are — emerge from a set configuration of various unconscious drives and their relation to one another.¹²³ Leiter refers to this as the “Doctrine of Types,” and argues that while Nietzsche recognizes a wide range in variability in the way that these “type facts” manifest in different individuals, they

¹²¹ BGE §19

¹²² D §119; GM §I.13; GS §334

¹²³ BGE §6

represent a kind of *nature* which is relatively fixed.¹²⁴ Accompanying the expression of our unconscious drives is the conscious *affect of willing*, a feeling which, according to Leiter, Nietzsche holds to be epiphenomenal.¹²⁵

Obviously, dismantling the conventional distinction of self and agency has deterministic implications which threaten to render any moral prescriptions or judgments meaningless.¹²⁶ It follows from this model that acts that we may judge as volitional, and therefore susceptible to moral approval or reproach, would instead merely be the culmination of drives and impulses leading us to the moment of action — each instance of such being a symptom of our nature (type facts) rather than a deliberate choice. Even assuming specific kinds of actions to be more valuable than others, it is unusual to present values in moral terms if technically speaking, *we* are not “free” to accept or reject directives based on them. We might be tempted to abandon ethical interpretations and simply accept that a kind of determinism follows from Nietzsche’s critique of agency. However, Nietzsche makes clear that he is not engaged in the wholesale rejection of moral value itself, but its widely-accepted foundations, in the hope that he may establish more useful and enduring forms of morality, or at least lay the groundwork for a future discourse in which morality will be untethered to those foundations.¹²⁷ Nietzsche’s new moral grounding will face the ambitious task of reconciling the possibility of moral responsibility with two truths about agency that conventional morality fails to acknowledge: 1) agency is a specific kind of phenomenon with causal origins misplaced

¹²⁴ Leiter (2015(a), 125)

¹²⁵ *Ibid.* 128; BGE §19

¹²⁶ Leiter (2015(a), 124); BGE §21

¹²⁷ BGE §214, §228

from where we commonly think them to exist (i.e., in our conscious faculties of deliberation) and 2) actions are the consequence of some kind of fixed disposition, variably instantiated across different types of individuals.

Moral beliefs are necessarily normative, in that the overall moral system in which the beliefs participate presumes some actions to be right and others to be wrong. This is true whether we follow the rules or not; if we subscribe to a moral view, then we accept some standards by which our actions are evaluated, and implicit pressures toward actions of particular kinds. These conditions are present in Nietzsche's moral philosophy even though his accounts of agency appear to allow little room for personal moral responsibility by conventional standards. For Nietzsche, normativity is a natural consequence of agency itself. This bears further explanation.

Given the inscrutability of the processes that drive our actions, under what conditions can we claim to have actual agency (a minimal sufficient condition for moral responsibility)? Nietzsche's concern is that when we consider the unconscious mechanisms and drives involved, much of what we take to be "active" in agency is actually passive, to some significant degree. Still, in order to make moral judgments, we must be able to make some distinction between activity and passivity in our actions. According to Paul Katsafanas, to begin with, we can distinguish *active* agency as actions performed in a state of equilibrium; in other words, agential actions are those that an agent approves of, with the condition that this approval would not be affected by any additional information concerning the actual cause of those actions. To put it another way, if we were to learn upon further review that our reasons for acting are not what we took them to be, but instead the result of hidden drives or undisclosed motives that are at

odds with our reasoning, then we would not consider our role in acting to be *active*.¹²⁸ This condition accounts for the possibility of circumstances in which we may be wrong about our presumed motivations for acting. The example that Katsafanas introduces is that of an anorexic patient, who deems it completely rational to avoid eating, but would likely not feel the same way about this justification were their decisions not influenced by the disorder.¹²⁹ This example represents (from the outside at least) a rather obvious case of disequilibrium, but we must remember that Nietzsche has consistently argued that much more of the reasoning grounding our active agency than we realize lacks the criterion of equilibrium.

Introspection into our motivations is the first step in determining if we are in a state of equilibrium, but there is a complication; we may not know if our introspective conclusions are themselves in disequilibrium.¹³⁰ To continue to dissect the layers of background motives in this way appears to only lead us into a regress. Actually, this is not necessarily the case. As Katsafanas points out, while the actual drives and motivations in question are technically out of the range of our introspection, we must remember that all such considerations are available to us in abstraction, as counter-factual models of agents and actions.¹³¹ Once again, remember that this is Nietzsche's domain of moral discourse — thought experiments which highlight the neglected corners of human psychology. While it is true that we may never reach that bedrock of causal priority, through counter-factuals, we can assess the conditions which would lead an individual to

¹²⁸ Katsafanas (2013, 110)

¹²⁹ Katsafanas (2013, 135)

¹³⁰ *Ibid.* 141

¹³¹ *Ibid.* 142

be in a state of disequilibrium, and then provisionally compare these conditions to our own presumed motives. The anticipation of “additional information” that could change our approval of our actions is an additional hypothetical safeguard against disequilibrium.

The concept of equilibrium provides a more precise means to distinguish between active and passive acts, thus satisfying our minimal condition for personal moral responsibility, but further steps are required in order to develop moral claims concerning what kinds of acts we *should* approve of, and why. According to Katsafanas, action has what he refers to as *constitutive aims*. We have seen that the first of these constitutive aims is equilibrium; in order for our actions to qualify as action at all, we must approve of our actions. The second of these constitutive aims is the *will to power*.¹³² The will to power is that principal drive by which all beings are motivated to pursue and overcome resistances — a drive which, according to Nietzsche, is all too often constrained in conventional Western moral systems.¹³³

2.4 Exemplars and the “Normative” Value of Will to Power

If, as Katsafanas claims, the will to power is a constitutive aim, it endows actions with a kind of “built-in” normative value. While there are always concealed and indeterminate drives motivating actions — often in conflict with one another — we can reliably assume the most primary of these to be the agent’s will to power. In addition to clarifying questions of motivation, we can, as Nietzsche consistently does, employ our understanding of this drive as a standard of moral evaluation, with a few caveats. Will to power can function as a normative principle because for agents, as a drive which is prior

¹³² Katsafanas (2013, 146)

¹³³ BGE §259

to all other drives, there is value in performing actions that successfully and consistently satisfy it. We might say that insofar as we approve of acting at all, we *should* perform actions which express the will to power.¹³⁴ To summarize the normative argument on Nietzsche's model of agency: 1) we are fundamentally constituted in such a way as to primarily will power, 2) because of this constitution of the will, and the way that the will to power is expressed, we are beings that generally thrive in circumstances that allow for the effective expression of the will to power through meeting and overcoming resistances, and because of these two claims, 3) action is valuable to the degree that it succeeds at its aim of willing power (by meeting and overcoming challenges). Similar to Kant's strategy in formulating the *Categorical Imperative*, Nietzsche has determined the *Will to Power* as a universal ground to agency, though one which is more resilient to his critiques, based upon what is actually accessible to our investigation – our actions.

Because of Nietzsche's particular view of power as the principle constitutive aim of our drives, and because of the way that the affect associated with the will to power is generated, we might think that we have an ironclad response to Silenus' warning of a human existence which is maximally and inescapably challenging. This is true to a certain extent, in that the level of possible resistances which such an existence provides is presumed to be inexhaustible. In other words, a life of suffering is a dynamic and unending flow of resistances for us to confront. However, the will to power is expressed in *overcoming* resistances, rather than being perpetually overcome by them.¹³⁵ We experience the affect associated with the will to power when we meet resistance and

¹³⁴ Katsafanas (2013, 155)

¹³⁵ Katsafanas (2013, 184)

overcome it, but to be constantly thwarted by resistances which we cannot overcome does not give us this feeling; instead, we feel the negative *affect* of being under the control of pressures beyond the boundaries of our will.¹³⁶ The ability to meet and overcome challenges and therefore successfully express the will to power varies according to the capacities of the individual, and so does the level to which the pessimistic view of Silenus can be embraced authentically, independently of the Apollonian veil of placating existential narratives.

Nietzsche's account of Silenus embodies the most challenging form of the pessimism which any life-affirming philosophy must confront: life is inherently without meaning or value, and suffering is inevitable. This view aligns with one which is heavily influential to Nietzsche's own — that of Schopenhauer — though the two philosophers arrive at significantly different conclusions about how we should act under these conditions. The mythical Silenus very clearly argues that there are only two possible responses to our conditions of existence: antinatalism and suicide, and the first of these is already unavailable to us. Nietzsche introduces the Silenus myth to represent a third option: avoidance of the inescapable and harmful truth — better perhaps that Midas had never demanded to uncover it in the first place. But assuming that this Pandora's box is now already opened, how may we successfully distract ourselves from the truth?

According to Schopenhauer, through artistic representation, we are afforded a kind of necessary but temporary “escape” from our immersion in this unending experience of adversity.¹³⁷ But Nietzsche offers a further pair of alternatives to the

¹³⁶ Leiter (2015(a), 122); BGE §19

¹³⁷Schopenhauer (2006); Reginster (2014, 19)

challenge of Silenus. While Nietzsche shares Schopenhauer's acceptance of the ubiquity of suffering, he diverges from Schopenhauer's prescription of the will's response to it.¹³⁸ For Nietzsche, the value of tragedy is not that our attention to and engagement with the world is diverted by a beautiful representation, but that an honest representation will illuminate or emphasize the actual beauty of an existence rife with suffering and tragedy. For those individuals who have developed the capacity to face the challenges that such a life represents, the experience of perpetual hardship constitutes an ongoing source of potential resistances to overcome. Alternately, for those who lack this capacity to such a degree, tempering the *Dionysian* truths in *Apollonian* narratives allows for intermediary stages of resistance and overcoming.¹³⁹ This is not the same course as the diversion Schopenhauer's aesthetic offers; for Nietzsche, life does not merely inspire art as a distraction, but life becomes art, whatever perspective we adopt as its voice.¹⁴⁰

Nietzsche's model of the constitutive aims of agency as the ground of value raises a troubling concern: while the will to power must express itself through perpetually meeting and overcoming resistance, this model does not distinguish a particular type of overcoming, opening the door to valuing behavior which we may determine to be egoistic, cruel, and even tyrannical by what we presume to be any reasonable standards. In fact, we find that Nietzsche often explicitly praises and defends these properties of character.¹⁴¹ However, Katsafanas points out that while willing power is valuable to us because it is constitutive of agency, and the successful expression of power is valuable on

¹³⁸ Reginster (2014, 21)

¹³⁹ BT §18

¹⁴⁰ Recall that Nietzsche holds any account of the world to be necessarily perspectival (Nehamas 1985).

¹⁴¹ GM §II.6; BGE §225

this basis alone, it is not the only value by which we judge our actions.¹⁴² The drive to express power may cohere within a framework of other values which include drives toward selfishness and cruelty. It is important to understand that Nietzsche does not praise these values except in relation to their extrinsic value in facilitating our successful expression of the will to power.

Much of Nietzsche's praise of cruelty as cultural practice is presented in one of two ways: either as reflection on what we might refer to as the *bygone* nobility of more violent cultures of the past, or as affirmation of the free expression of a component of human nature that we may attempt to sublimate but can never fully overcome.¹⁴³ When Nietzsche urges us to consider the value of cruelty, he observes that whatever we may make of it, we feel powerful when we are cruel toward another. Such instances of cruelty also offer opportunities for those enduring torment to will power through resisting and overcoming their torment.¹⁴⁴ Similarly, there is some benefit in cruelty which is self-directed — to willfully place oneself into the path of suffering which could otherwise have been avoided in order to meet a greater range of challenges.¹⁴⁵ These occasions for *constructive* cruelty stand in contrast with a kind of cruelty Nietzsche criticizes, which is the way that forms of modern culture inflict cruelty on individuals by *limiting* their potential to successfully will power.¹⁴⁶ Suffering for its own sake has no intrinsic value,

¹⁴² Katsafanas (2013, 148-149)

¹⁴³ GS §23, §362; BGE §225; GM §II.6-7; In *Ecce Homo*, Nietzsche refers to cruelty as "...one of the most ancient and basic substrata of culture that simply cannot be imagined away" (EH:GM).

¹⁴⁴ Parmer (2017, 402); GM §II.6

¹⁴⁵ GS §266

¹⁴⁶ BGE §229

and this includes the infliction of suffering with no associated possibility of overcoming, such as that imposed by the ascetic, or other “life-denying” philosophies.¹⁴⁷

When Nietzsche urges us to reconsider our accustomed rejection of selfishness, barbarity, and cruelty, it is because these values represent one effective pathway to fulfilling the primary constitutive aim of our actions — power. However, there are others which are better suited to the conditions of our existence. Even though there is extrinsic value in suffering, there is nothing to be gained by intentionally creating more suffering for others, as life provides ample opportunities for suffering already. It is improbable that we can completely discount the extrinsic value of suffering, but we can understand that in the long run, it may lead to greater internal conflict among our other values and drives, thus actually interfering with the successful expression of the will to power. It is preferable to seek out resistances through self-cultivation and creativity, and in this way eliminate, or at least minimize any internal conflict between our various drives and values. Katsafanas notes that the exemplars Nietzsche holds in the highest regard are not the violent tyrants engaged in cruel subjugation over others, but artists and intellectuals, willing power by setting their energies upon creative tasks.¹⁴⁸ This is reasonable if we consider Nietzsche’s encouragement for us to live a kind of life that is valuable on primarily aesthetic terms. Whatever forms such existence may take, because of the priority of the aesthetic, it makes sense that artists would occupy a privileged status in Nietzsche’s articulation of exemplar narratives.

¹⁴⁷ BT §7; GM §III.11

¹⁴⁸ Katsafanas (2013, 185)

2.5 Defining a Nietzschean Model of Virtue

Because the will to power motivates all of our actions, we have a reason to resolve conflicting values which might restrict or inhibit the success of actions which most effectively express power. Agency consists of an aggregate of processes and drives infinitely more complex than we can ever discern in their totality, so it is unlikely that we can ever completely eliminate internal conflict, but we can take steps to minimize it. For Nietzsche, the worst kind of internal conflict arises when we adopt values which express themselves in ways which fundamentally conflict with the will to power. While the will to power is a necessary, immutable feature of agency, presumably, we have some control over how we develop other values. Nietzsche's moral philosophy amounts to a re-examination of the value of our values, with power as the standard by which we conduct this undertaking.¹⁴⁹ This strategy is crucial, because through the course of things we may adopt or enculturate values which impede the will to power, and thereby worsen the internal conflict of our drives. As Katsafanas' argues, simply by virtue of these facts about the nature of agency, power functions as a prescriptive criterion.

At first glance, the course corrections that Nietzsche prescribes may seem like a simple series of choices to change behavior in order to produce the kinds of actions that cohere best with the will to power. The problem with this, as Leiter suggests, returns us to the questionable presumption of the identity of self and agency. We think that we intuit a situation which calls for us to *will* a correction in our behavior, and then subsequently this correction takes place in accord with the causal influence of the will. However, according to Leiter, for Nietzsche, this relationship is only epiphenomenal; the

¹⁴⁹ Katsafanas (2013, 188)

disposition to enact specific actions — including those which constitute amendments of prior dispositions and actions — are actually only a result of stable “type facts” over which we have no actual control.¹⁵⁰ In the same way that sets of inscrutable drives motivate actions, others may motivate changes in those actions, independently of our perceived decision to make adjustments in accordance with any moral criteria. According to Leiter, “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.”¹⁵¹ This means that while there may be a moral prescription that we endorse in equilibrium, whether we follow this prescription or not is in no way a result of our conscious willing, but the outcome of a “contest” of deterministic impulses.

If this is the case, can we actually change our values? Or is the capacity to change them already a fixed element in our nature — a disposition to self-cultivate, to realize a certain kind of self? At times, Nietzsche does appear to hold an entirely deterministic view of agency.¹⁵² This is problematic because determinism is fundamentally incompatible with the idea of moral responsibility. However, once again we face the challenge of reconciling these apparently deterministic views with those which explicitly prescribe self-cultivation and which appear to acknowledge the possibility that our drives can in fact change.¹⁵³ A belief in the possibility of self-cultivation entails the idea that not only can we change and evolve, but that we have some influence over this change, and are able to affect the course of events that determine what kind of person we become. The

¹⁵⁰ Leiter (2015(a), 129)

¹⁵¹ Ibid. 133. Leiter refers to an aphorism in *Beyond Good and Evil*: “The will to overcome an affect is ultimately only the will of another, or of several other, affects.” (§117)

¹⁵² GS §127; BGE §21.

¹⁵³ D §109

reason that these views do not stand in contradiction is that Nietzsche is not a skeptic with regards to the possibility of agency and self-cultivation, but with regards to something more basic — our conventional understanding of causation.¹⁵⁴ In the case of agency, the error occurs when we presume that we cause our actions through a consciously directed effort of will.

It may not be clear how this distinction avoids the complications of determinism. Recall that in the “extra-moral” phase, Nietzsche directs us to abandon intention as the criterion of value, and instead concern ourselves with what is *unintentional* in the act.¹⁵⁵ If we are mistaken in ascribing causal efficacy to the phenomena of intention, then it makes sense that Nietzsche would advise us to detach ourselves from this error. It is more difficult to understand what Nietzsche actually means in prioritizing the unintentional. Aaron Ridley interprets this passage to mean that the “causal” will and subsequent action are not known to us as discrete phenomena, but are merely conventional differentiations of a single “mutually constitutive” event — the act itself.¹⁵⁶ In *On the Genealogy of Morals*, Nietzsche explains: “[...]there is no “being” behind doing, effecting, becoming; “the doer” is merely a fiction added to the deed – the deed is everything.”¹⁵⁷ In more general terms, it makes no sense to speak of a cause without an effect and so a will without an action has not actually willed anything. Willing is an event which comes into our awareness only at the time that the “willed” action occurs. If, as Leiter concludes, the association between willing and action is epiphenomenal, this may account for the

¹⁵⁴ BGE §21

¹⁵⁵ *Ibid.* §32

¹⁵⁶ Ridley (2018, 17)

¹⁵⁷ GM §1.13

proposition that inaccessible “type facts” lead us to act in specific ways.¹⁵⁸ However, if the relationship between the will and action is *only* epiphenomenal, then how do we have any influence over the type facts which are the true causes of our actions, as Nietzsche’s (and in fact, any) moral prescriptions would require? How can we effect a legitimate moral change on our underlying properties of character such that we may create consistent patterns of responses to moral challenges? In other words, we appear to have circled back on skepticism about the possibility of agency.

The set of type facts that dictate our actions are what Lester Hunt refers to as our “deep character.” In order to change our actions, we would require the ability to adjust either the deep character which leads to our actions, or the actions themselves. Hunt reasons that the first option is impossible, and the second is “not ethically relevant,” as it represents a surface-level conformity which is only practical, rather than a true modification of individual moral properties.¹⁵⁹ There is very often a “practical” concurrence between actions and what we take to be the intentions that motivate them. If this were not the case, then we would likely find it difficult to make judgments about whether our actions were undertaken in equilibrium. The problem is that our only access to deep character is not “deep” within the hidden psychology of the agent, but outward, in actions.¹⁶⁰ So if we accept the idea of one-way causality between deep character and surface behavior, then morally relevant change is simply not possible.

However, Nietzsche does believe in the possibility of our behavior influencing changes in the drives which comprise deep character. In *Daybreak*, he outlines specific

¹⁵⁸ Leiter (2015(a), 129)

¹⁵⁹ Hunt (1993, 17); WTP §394

¹⁶⁰ Ridley (2018, 79)

ways by which we can habituate responses through repetitive actions, and in this way weaken, or even altogether eliminate the influence of drives over our actions.¹⁶¹ This model demonstrates clear parallels to the mechanism of moral development under classical virtue ethics. To clarify, Aristotle believes that we are capable of developing moral virtues through the repetition of specific kinds of actions in specific circumstances, and that these actions come to us more naturally in what we determine to be similar circumstances.¹⁶² Some actions may become a kinds of default responses, as habituated expressions of an agent’s character, or deep character (drives and impulses). Once such a characteristic has been sufficiently cultivated, one may reach a state in which the individual is not in conscious negotiation of each action in order to cultivate or maintain the desired kind of conduct. Instead, good action “flows” from a good kind of character, and eventually, becomes internalized as a disposition of that character.¹⁶³ If an individual has conditioned themselves toward certain patterns of behavior, then there is some component of the new behavior which is generated by the disposition toward that kind of action, and not the “surface” decision to act a certain way. This provides reasonable justification for the claim that, in some sense at least, a virtue is a point of convergence between doer and deed, rather than a presumption of the former as the “substratum” for the latter, which would entail unsound metaphysical speculation.¹⁶⁴ In other words, the expression of a virtue is an event in which a specific kind of character performs a specific kind of action. If the agent *lacks* the disposition toward this virtue, they still may hold the

¹⁶¹ D §109

¹⁶² NE 1103a25

¹⁶³ NE 1103a15

¹⁶⁴ GM §1.13

capacity for its development, through manipulation of the mechanisms that comprise deep character.

Virtue functions as a more stable causal grounding for morality than intention because it does not look to the inscrutable “why” of an action, but to the indisputable “what” in its expression. When Nietzsche collapses agency *into* action, it suggests that action consistently arises dispositions, the cumulative result of drives which comprise deep character. While we have no direct access to the components of deep character, the important moral part of a virtue is in its expression, in what is accessible to us — the deed. Virtue ethics works in this way because drives are dispositional, not volitional. They are events in which specific kinds of characters perform specific kinds of actions. The standard of *good* action is then the act undertaken by a *good* agent – what we would refer to as an *exemplar*. From outward knowledge of an agent’s patterns of behavior, we may derive indicators of habituated dispositions toward certain kinds of actions (virtues or vices). We see no “doer behind the deed,” but a doer *in* the deed. Contingent upon what traits of character we value, we develop standards of the exemplar type — the individual who appears to have internalized a broad range of virtuous dispositions, which we observe as an overall stable tendency to demonstrate valued behaviors in practice.

In virtue ethics then, the prescriptive component of good action is not determined by rigidly normalized sets of rules and principles, but through the emulation of worthy ideals. These ideals are actions which reflect the virtuous character of exemplars.¹⁶⁵ This standard of the good comes to represent a flexible body of value norms, as what is considered a virtuous character is contingent upon the shifting values of individuals and

¹⁶⁵ Brobjer (2003, 67)

groups. For both Nietzsche and Aristotle this flexibility is a fundamental feature of sound ethical commitments. Virtues are not universal maxims but are customs which evolve through refinements in a culture's moral reasoning, and its attunement to the particularity of circumstances which merit the intervention of this reasoning.

2.6 In Defense of the Virtue Theory Interpretation

If virtue ethics functions in the way that I have defended, it can accommodate the necessarily shifting and contingent nature of viable value sets in relation to the dynamic background conditions of our existence, and account for the problem of agency through the reorientation of value criteria into actions themselves. These features align with Nietzsche's ethical comments, but there are also differences between most familiar forms of virtue ethics and many of the values that Nietzsche embraces. As I have argued, in order to adequately defend Nietzsche as a virtue theorist, some additional distinctions must be established in order to account for notable differences between Nietzsche's specific virtues and those of our most familiar counterpart – in this case, the virtues of classical eudaemonist ethics.

We may anticipate the objection that the virtue theory interpretation of Nietzsche's ethics fails on the grounds that the resemblances to the familiar Aristotelian model are cosmetic at best, and that there is nothing in Nietzsche's work that resembles what can be called a virtue without straining the concept beyond conceptual coherence. One such critic, Jessica Berry, argues that Nietzsche's philosophy fails to meet reasonable criteria which would clearly establish Nietzsche as a virtue theorist. Berry refers to the dual conditions outlined by Christine Swanton which state that any virtue ethics must both identify what qualifies an action as "right," and then identify what traits

qualify as “virtues.”¹⁶⁶ According to Berry Nietzsche does not address either of these conditions at any point in his writings, leading Swanton to exaggerate textual support for the conclusion that Nietzsche is a kind of virtue theorist.¹⁶⁷ In the case of the first condition, Berry claims that Swanton falsely attributes to Nietzsche an evaluation of the goodness of action based upon virtuous *intent*. At issue is a problematic interpretation of this passage from *On the Genealogy of Morals*: “The popular mind in fact doubles the deed; when it sees the lightning flash, it is the deed of a deed...”.¹⁶⁸ Swanton appears to argue that in this passage, Nietzsche maintains the distinction between doer and deed, rather than rejects it. By this reasoning, action would be that which is “caused” by the prior phenomenon of intention. The point of this interpretation is to then hold, in familiar Kantian terms, the good intention as the moral criterion grounding the system of virtue. A character which *intends* good action is an exemplar, by the standards of this interpretation. If this is Swanton’s primary argument in support of a Nietzschean Virtue Theory, it is defective, as it either overlooks or confuses Nietzsche’s skepticism of the causal relationship between intention and action, as well as his disavowal of the *moral* value of intention.

Berry responds to Swanton’s second criteria: that a virtue theory must establish which traits are virtues. According to Berry, Nietzsche’s value claims do not function as virtues, nor does he intend them to, and so the virtue theory interpretation fails on this criterion as well.¹⁶⁹ Berry argues that virtue proponents such as Swanton have conflated

¹⁶⁶ Swanton (1998, 29)

¹⁶⁷ Berry (2015, 370)

¹⁶⁸ GM §1.13

¹⁶⁹ Berry (2015, 374)

Nietzsche's argument for the transcendence of outmoded systems with an argument for the advancement of a new system with which to replace them. There are problems with both Berry's and Swanton's arguments. Berry's criticism relies upon a constrained definition of the concept of virtue in order to reject Nietzsche as a virtue theorist, while Swanton's argument appeals to a flawed interpretation of Nietzsche's conception of agency. Though Berry succeeds at refuting Swanton's initial argument on these grounds, if it can be established that Nietzsche both provides an account of what he considers to be good actions, and an account of what traits are successful at bringing about these good actions, then on the criteria accepted by both Swanton and Berry, the pro-virtue theory argument still succeeds.

Berry concedes that Nietzsche presents prescriptive arguments to the extent that he encourages us to reach a point where we no longer require deference to systems that guide our actions, but she stops short at accepting that the values he advances throughout his writings can correctly be termed virtues. For Berry, the misguided urge to overstep and attribute positive moral content to Nietzsche without corresponding textual support overlooks what is most compelling about Nietzsche's work: the success of the critiques themselves.¹⁷⁰ It is true that Nietzsche's work is philosophically significant even if he only intends to dismantle foundations and does not replace them with his own moral propositions. If this were the case, Nietzsche's positive claims, his judgments, his praise of humans who fulfill the kind of existential narratives that he juxtaposes with the targets of his critiques, would all amount to kinds of literary exercises — ways to illustrate what he disvalues in systemic structures by illustrating their antitheses. For one thing, it is

¹⁷⁰ Berry (2015, 374)

unclear how this objection constitutes a necessary conceptual distinction, given Nietzsche's view of the relationship between art and life — a beautiful representation of life can function as an implicit call to align one's ways of living with the aesthetic ideal. Further, in his rejection of conventional morality, Nietzsche often explicitly encourages readers to cultivate the specific properties of character which would contribute to the further "good" of *extra-moral* flourishing. Berry's objection over referring to these properties of character as virtues would require a somewhat narrow restriction on the terminology.

There is a reason that the values which Nietzsche advances throughout his writings bear striking differences from familiar formulations of virtue ethics. Given the fundamental dissimilarity between Nietzsche's worldview and those of other virtue theorists, we should expect Nietzsche to endorse distinctly different kinds of virtues as different responses to kinds of moral challenges. In clarifying this distinction, Robert Solomon characterizes Aristotle's virtues as a response to the warlike Homeric tradition which preceded him, in which we derive the standard of good action by emulating the behavior of "heroes."¹⁷¹ Nietzsche's virtues, on the other hand, would reasonably emerge as a response to a post-enlightenment world view, in which a general awareness that the world is fundamentally incompatible with flourishing in the classical sense has grown more salient.¹⁷² Alasdair MacIntyre argues that the rupture between these virtue traditions is more complicated than we may presume, and that this has had a far-reaching impact upon the failures of contemporary ethics. MacIntyre views Nietzsche's morality not as a

¹⁷¹ Solomon (1985, 254)

¹⁷² BT §3; HH §517

repudiation of Aristotelian ethics, but as a response to the failure of the enlightenment tradition which broke with Aristotelian ethics.¹⁷³ While Aristotle’s virtues remain fixed to a “bygone” cultural context, MacIntyre argues that our values are still grounded in the “ghosts” of Homeric and Aristotelian values. While MacIntyre ultimately believes (incorrectly) that Nietzsche’s untethering of grounded value leads inevitably to relativism, his argument prompts us to consider that if the departure of enlightenment ethics had never taken place, would we still need to reject the Aristotelian virtue tradition?

According to Nietzsche, the answer is yes. Nietzsche provides ample argument that a rejection of traditional virtues is necessary beyond enlightenment ethical systems (and their subsequent failures). According to Nietzsche, those virtues which we have traditionally cultivated – those of diligence, obedience, chastity, piety, and justice – are those which aim to preserve the interests of societal agenda at great harm to the individual.¹⁷⁴ Nietzsche argues that our elevation of these values amounts to “...the praise of drives which deprive a human being of his noblest selfishness and of the strength for the highest form of self-protection.” In other words, these virtues conflict with the will to power. Nietzsche explicitly warns of the “leveling” pressures of a society which would effectively make us progressively more *mediocre* in the utilitarian interest of maximizing the human good in the universal sense, and at the cost of abandoning what is exceptional about humanity in its various particularity.¹⁷⁵

¹⁷³ MacIntyre (2015, 257)

¹⁷⁴ GS §I.21

¹⁷⁵ GM §I.12

Nietzsche's rejection of these long-standing virtues does not constitute an attack on the possibility of virtue itself, and he gives us some idea what kinds of virtues he would have us adopt, and how these virtues function in relation to our other drives and dispositions.¹⁷⁶ Like more traditional conceptions of virtue, Nietzsche's virtues reflect a cultivated disposition to enact values. Nietzsche assigns provisional value to properties which result in the transcendence of normalized systems, and those which reinforce the disposition toward properties of character which weaken our reliance on "herd" values. In other words, the move toward breaking out of the cycle of cultural conditioning is facilitated by a different kind of virtue, and like any other virtue, it can be cultivated under the right conditions, and through the right kind of practice.

What we can concede to Berry is that she has successfully refuted any argument for Nietzsche supporting an "Aristotelian" form of virtue ethics in his writings, but this is not the point that most proponents of a Nietzschean virtue ethic are advancing. Instead, there is strong support for the argument that Nietzsche develops a model of value formation which both emphasizes the active expression of some kinds of properly developed character and intends some further "good" state of affairs for us, to the extent that this is possible, given the circumstances to which we are bound. Value structures which are primarily focused on maintaining social order and preserving the status quo actually limit our capacity to achieve any sustainable level of flourishing. They do this in two ways. First, by dismantling meaningful narratives by which we aesthetically mitigate intolerable suffering, and second, by imposing new narratives of control, limiting our potential to effectively meet and overcome resistances. A Nietzschean virtue would be a

¹⁷⁶ BGE §214

property associated with the unwillingness to confine oneself to systems of value, and our values would be those of valuelessness, our system “system-lessness.” Nietzsche’s standard of good value involves the achievement of circumstances in which we are not bound to structures which harmfully inhibit our capacity to will power.

Both Christine Daigle and Mathew Dennis offer promising ways to flesh out Swanton’s two requisite conditions of a virtue theory.¹⁷⁷ According to Christine Daigle, in response to the first condition, Nietzsche’s specific virtues are those dispositions which further one’s progress toward the ultimate transcendence of value structures. In response to the second condition, these dispositions include those of creativity, individuality, free thinking, and “trans-systemic” reasoning — in other words, the propensity to forge one’s own existential narrative.¹⁷⁸ Mathew Dennis introduces his own subtle refinement to this account. For Dennis, the “ethical ideal” which is the ultimate end of Nietzsche’s virtue ethics is “becoming oneself,” and he articulates the specific Nietzschean virtues as those which lead us to self-cultivation.¹⁷⁹ Once again, this conception of virtue is at odds with predominant moral narratives, but following from the conditions of our existence and necessary facts about agency, in Nietzsche’s virtues we see both an articulation of which kinds of actions are valuable for us, and what kinds of character traits we *should* cultivate. In other words, Nietzsche’s ethics satisfies the conditions of virtue ethics, on criteria accepted by both Swanton and her critic Berry.

¹⁷⁷ To restate Swanton’s thesis: virtue ethics must both identify 1) what qualifies an action as “right,” and 2) what traits qualify as “virtues.” (1998, 29)

¹⁷⁸ Daigle (2015, 406)

¹⁷⁹ Dennis (2019, 56)

2.7 Self-Cultivation, Overcoming, and the Human Narrative

The practical strategies of conventional virtue ethics align with Nietzsche's insistence that our moral progression has led us to this point that "the decisive value of an action lies precisely in what is unintentional in it."¹⁸⁰ When actions become habits, then we are less dependent upon navigating the problematic psychology of intention. However, conventional theorists such as Aristotle hold a much more optimistic view of the efficacy of our capacity for self-cultivation than does Nietzsche, most likely owing to more robust ontologies of self and agency. Ridley observes that Nietzsche's prioritization of self-cultivation can be summarized in two seemingly contradictory encouragements: to both become who we are, and to overcome who we are.¹⁸¹ If the foundation of agency — our deep character — consists in type facts that are largely inaccessible and *unchangeable*, then any encouragements to "become" this nature would strike us as irrelevant, because how could we do otherwise? Similarly, any instructions to "overcome" these type facts seem to be demanding the impossible.

In order to interpret this paradox, it only makes sense that Nietzsche has two different kinds of deep character in mind, and that he encourages us to instantiate or "become" the type facts in the case of the first, and to overcome them in the second sense. The first kind of type facts fall into the category of "capacities," and include those features of our character which are fixed, and unchangeable. The second kind of type facts we can assume to be "dispositions," and these constitute habituated and *malleable* patterns of drives which lead to regular kinds of behaviors, such as the virtues discussed

¹⁸⁰ BGE §32

¹⁸¹ Ridley (2018, 168); GS §270; Z §1: "On the Three Metamorphoses"

above. To restate the virtue theory argument: it is reasonable that through practice and self-cultivation, we may effect change over our dispositions in order to develop new ones or weaken old ones. In interpreting the call to “overcome” the self, we might conclude that Nietzsche wishes us to enrich our state of being by acquiring new and better virtues. This is no doubt true, but it is only a part of Nietzsche’s reasoning. In *Thus Spoke Zarathustra*, Nietzsche writes: “Ten times a day you must overcome yourself, that makes for a good weariness and is poppy for the soul” and then “Ten times you must reconcile yourself again with yourself, for overcoming causes bitterness and the unreconciled sleep badly.”¹⁸² This implies that there are two effects of the same process of overcoming. This is possible because these effects occur in two ways, at two different levels of our deep character. When we direct our energies to self-overcoming, and to cultivating our dispositions, this is beneficial to us, because we achieve access to a better kind of narrative of self, and we may improve the state of our drives to bring them into a more beneficial harmony with that which cannot be changed: the type facts which comprise our capacities.

While we can apply this distinction to argue the possibility and advantage of self-overcoming, it may not be initially clear how it accommodates the instruction toward self-becoming. In a famous aphorism in *The Gay Science*, Nietzsche writes: “*What does your conscience say?* — ‘You should become who you are.’”¹⁸³ Once again, the most reasonable interpretation is that in passages such as this, Nietzsche is urging us to engage with our deep character, only this instruction is an appeal to one’s capacities — the deep

¹⁸² Z §1.3

¹⁸³ GS §270

character which is a fundamental and static feature of who we are. Why is this a significant kind of claim — to become what we already are? And why does this constitute such an important part of the overall process of developing the narrative of self? In *Ecce Homo*, Nietzsche expands on this recurrent theme: “To become what one is, one must not have the faintest notion *what* one is.”¹⁸⁴ The reason that becoming what we are is an important ethical problem rather than a simple redundancy is because we might be wrong, or worse: misguided about what we are. In fact, to restate Nietzsche’s view, we cannot know our deep character in any meaningful or reliable way. All that we know of our deep character is that as agents we are fundamentally constituted with both the capacity and disposition toward the will to power. How effectively we are able to will power is a matter of whether the sum of our other values and drives cohere or conflict with this fundamental aim.

The fact that we *cannot* know *what* we are is an important clue to interpreting Nietzsche’s claim that a necessary condition of becoming is not actually knowing what we are. To begin with, if we believe to have knowledge of the inaccessible drives and motives which comprise deep character, we are likely to be wrong. If we hold a narrative of deep character as a model of our essential being, then it is a short step to letting this narrative dictate the course of our choices, actions, and any power we may have over the development of our dispositions. To relinquish the course of our potential development to a flawed idea of our limitations is to accept unnecessary constraints on what we are capable of becoming. Here is the heart of the problem — we may be misled into accepting such a defective narrative in the interest of purposeful restriction of the range

¹⁸⁴ EH §11.9

of our capacities, in order to hold our development into a prescribed range, an egalitarian mean. While some forms of egalitarianism are beneficial, the kinds which restrict the growth of those with exceptional capacities to bring them into balance with a range of capacities which is universal to all individuals will effectively eliminate the expression of exceptional character.¹⁸⁵ Once again, Nietzsche's concern lies with who appropriates control of the narrative of human agency, and to what end they set the terms of what it means to fall into a prescribed and range of human dispositions and capacities. Recall that simply by virtue of the conditions of authentic agency, there is normative value in avoiding circumstances which inhibit our potentiality to become who we really are – our maximum manifestation of drives in harmony with the will to power.¹⁸⁶ Because some individuals lack the capacity to will and overcome resistances at the same level of others, given the opportunity, these individuals could seek to alter the standards of what constitutes both normal and exemplary capacities, and to determine what specific virtues we should hold and in what degrees.¹⁸⁷

In the last chapter, I discussed the normative value of exemplar narratives that function as a kind of ideal range of human virtues. Lacking access to any metaphysically necessary foundation of selfhood, we only have access to *perspectives* of an ontology — our narrative representations of what it means to be human, including what it means to become an exemplar type. For Nietzsche, merely by virtue of being agents, our constitutive aim of willing power entails the ongoing pursuit of resistances to overcome. The ultimate manifestation of perpetual overcoming is of the self. We continuously push

¹⁸⁵ BGE §228; GS §116

¹⁸⁶ Katsafanas (2013, 146)

¹⁸⁷ *Ibid.* 220

the boundaries of this narrative, testing our capacities to overcome new levels of resistance, and to the extent that we have the power, we modify our dispositions to bring our values and our drives into coherence with better forms of the narrative. We may presume that a human being is free to become and overcome their narrative of self within the constraints of the collective type facts which comprise their capacities. But even this is not completely true. Dispositions place further contingent constraints upon how effectively we fulfill the ultimate potential of our capacities. Determinations about the kinds of dispositions that we cultivate are influenced by our values. If we allow ourselves to internalize the values of others who would impose their will over us, as a mechanism of control, then our capacity to “become who we are” is compromised.

In the chapter that follows, I will explore specific ways that this anxiety over who controls the human narrative manifests itself in aesthetic representation. Each of the science fiction narratives that I will discuss builds upon a dichotomy of authenticity, represented in the difference between humanity and artificial versions of humanity. Importantly, as the history of science fiction cinema and literature began to approach this dichotomy in more sophisticated ways, ontological narratives began to reveal our intuition that the conventional approach to the dichotomy is flawed. In breaking down our presumptions of what it means to be human, to become human, and to instantiate a range of the exemplary virtue set of humanity, these narratives also address how our culturally imposed versions of the human narrative can limit our fundamental capacities through narratives of institutional control, othering, and abuse.

CHAPTER 3

SCIENCE FICTION FILM AND THE HUMAN NARRATIVE

3.1 Introduction: The Moral Value of Narrative

Nietzsche's criticisms of morality in the pejorative sense reduce to the same general argument: every principle which presumes an objective ground of truth in order to derive moral value already rests upon a series of myths about the world, ourselves, and the relationship between. Consequently, such systems cannot prescribe responses to ethical challenges with the rigid objectivity that they presume. Worse, the untenable myths such as those advanced by MPS are self-defeating — slipping points on the way to relativism and nihilism. While conventional moral skeptics may regard the “ungrounding” of truth correspondence as the negation of value, for Nietzsche, there is no actual epistemic advantage in dismissing the possibility of knowledge simply because we cannot access its ground as a necessity, nor in rejecting affirmative ethics because we cannot identify the origin or nature of agency. In the way that we order our experiences, truth and illusion are something that we synthesize in a world, and in a fundamental sense, to posit patterns in one's deeds is to implicitly give an account of agency. We think of consistent patterns of actions as reflecting some form of the *deep character* of agents. In forming existential narratives, we then impose value on certain ways of living above others based upon specific perspectives on the circumstances in which we exist. Our most valued of these perspectives express the character of our exemplars — the ways that agents should (and

do) act, accounting for a *mythologized* relationship between self, society, and world. In this way, we find answers to questions concerning what our ethical obligations are and how these obligations are justified. The challenge we face is in enculturating sustainable myths — for Nietzsche this means exemplar narratives that successfully embody aesthetic merit.

3.2 Narrative, Genre, and an Account of Virtue

Narrative is an important evaluative tool for us. In the same way that any account of the world is inescapably perspectival, narrative fiction can communicate value to us by limiting what we experience in the representation of the world into specific perspectives, and by guiding us to focus on specific features of its subjects' fictional lives.¹⁸⁸ In this way, popular narratives tend to exemplify moral perspectives by showing us a range of character types and actions, and through various techniques of form and style, by guiding us to evaluate some positively and others negatively. As Joseph Kupfer states, “Narrative arts have always emphasized the character of their characters [...]”¹⁸⁹ Because our judgments of moral character tend to be based upon consistent patterns of action, virtue is most easily understood if we have access to the history of an agent. Narratives present simulations — thought experiments — involving how fictional agents (characters) act in a variety of different situations (plot), and thereby reveal what may constitute worthy virtue sets for us.¹⁹⁰ Through narrative form, we have the opportunity to “rehearse” what the exemplary human life does and does not look like. Unburdened by the contingent

¹⁸⁸ Nehamas (1985, 55)

¹⁸⁹ Kupfer (2006, 341)

¹⁹⁰ *Ibid.* 345

constraints of existing within the realm of the familiar and the possible, details of our existence can be abstracted to test conventional ethical models and explore how a number of beliefs that we take to be essential to the moral narrative are actually ill-conceived or inconsistent with other values that are more fundamental to us. Kupfer elaborates further on the relationship between narrative art and ethics: “Audiences, elite and popular, are drawn to questions of virtue and vice perhaps because so much of our welfare depends on our own character and because we prosper and suffer at the hands of other people on account of their moral traits.”

In other words, knowing how to recognize different kinds of moral character in the absence of objectively grounded moral systems is so central to our realm of concern that virtue is thematically pervasive in our narratives. Themes and ideas that are matters of enduring concern for us are those which we are drawn to re-enact in various forms. A popular theory known as the “ritual performative theory” holds that we are motivated to re-enact the same kinds of themes and situations over and over again because of a deep psychological drive to ritualize enduring values through practice and repetition.¹⁹¹ When we consider kinds of narratives, we intuit that most can be classified into informal categories based on some criteria of similarity. We know these categories as *genres*. According to the ritual theory, genres are an extension of the process of symbolically maintaining things that we value through ritual performance.¹⁹²

While examples of narratives exploring moral themes may fall into a variety of genres, here I will argue that science fiction narratives are uniquely situated to ritualize

¹⁹¹ Bordwell (2019, 338)

¹⁹² *Ibid.*

our interests and concerns related to our formulations of virtue sets and the exemplar types that exhibit them. Exemplars in science fiction narratives often exhibit a range of virtues that diverge from the conventional sets, and into the realm of Nietzschean value: artificial beings *perform* the necessity to *create* a self in the absence of an objectively grounded *essence* of self.

Like other genres, the criteria that conventionally distinguish a narrative into the category of science fiction are broad. There are often elements which stand in some kind of noteworthy contrast with our *actual* world. These elements may include an emphasis on advanced technology, travel to alien settings, or significantly for this discussion, narratives dealing with robots, cyborgs, AI, clones, or other kinds of beings presented as “simulating” humanity in some way. Author Philip K. Dick argues that while any of these elements are commonly present in science fiction narratives, the distinguishing characteristics of science fiction are more general, over and above any of these thematic elements. According to Dick, the hallmark of a science fiction narrative is that it presents either an unusual or “eccentric” perspective of an ordinary setting, or an ordinary perspective of an unusual setting, and then introduces something novel — an idea — which initiates a change in this setting, impacting the lives and world of the characters. Dick notes that what distinguishes *good* science fiction from the commonplace or inferior is that the best narratives often do more than simply astonish audiences with the introduction of the outlandish; they posit answers to hypothetical questions about these novel ideas and extraordinary elements, and in the process, illuminate some unexplored aspects of our *actual* world.¹⁹³ In other words, science fiction stories can function as

¹⁹³ PKD (1995, 44-45)

counterfactual thought experiments, with some impossible or improbable element realized in the diegesis to highlight the effect of the novel upon the familiar.

3.3 The Human Narrative

In presenting a narrative which is bound up in a sense of wonder with a world that is both fundamentally like and unlike our own world in some important ways, most contemporary works of science fiction offer us either a perspective of something imperfect about the world and a way it might ideally be otherwise (the “Utopian” perspective), or a dark glimpse of where we may be headed if we are unable to alter the trajectory of our society (the “dystopian” perspective). Through the lens of these perspectives, science fiction narratives prompt us to scrutinize various conventional beliefs concerning what it means to be the kind of beings that we are, and to value different kinds of actions on this basis. Because of this, those ontological narratives of human types are well-suited thought experiments in ethics. These specific kinds of science fiction narratives effectively illustrate the arguments developed in the prior sections – specifically, that our moral values are most stable if grounded in aesthetic models, and that among the most fundamental of these aesthetic models is the narrative of what it actually means to be a human being, in some kind of meaningful sense.

Through the narrative conceptualization of a human life, there is an exemplar model which situates us in relation to an ideal moral standard to identify with, one that we may also value as aesthetically pleasing. The human protagonists of science fiction may not always be “exemplars” in the traditional sense, but they are nevertheless hypothetical moral surrogates, exploring what is actually at stake for us in the often

arbitrary and misguided process by which we distinguish what qualifies as authentic or artificial expressions of virtue and exemplary “human-ness,” in a setting in which, strictly speaking, biologically human beings may not be the only beings capable of moral reasoning and responsibility.

An effective way to understand what a functional representation of the human exemplar would look like is to contrast it with what it is clearly not, through a series of counterexamples. Assume that an authentic human being is understood to possess a minimum threshold of some essential characteristics. The “counterfeit” human being of science fiction narratives will often function in one of two ways. First, the “simulated” version of the human being will possess an incomplete set of these characteristics, thereby emphasizing the authenticity of its “real” human counterpart by comparison. This essential *lack* may present the artificial being as a threat to the world of the authentic beings in the narrative. Alternately, the artificial, or “simulated” human may be portrayed as inauthentic based upon some criteria which we discover to be insignificant or arbitrary, and in this way, the narrative criticizes standing normalized criteria by which we would account for authenticity in humanity.

When we quantify what it means to be human, we generally want to employ criteria beyond simple organism or mechanism. We usually have in mind a specific range of characteristics, analogous to a virtue set — something moral, something which is revealed in one’s deeds, and not in one’s physical makeup. When we consider narratives which instantiate kinds of exemplars, the human virtue set may serve as a narratively constructed “essence” of the exemplary human type. Where science may offer at least some persuasive arguments to the contrary, we aspire to narratives that hold the essence

of humanity to be something more noble than an accident of biology. Though this is optimism on our parts, we see the right kinds of deeds as exemplifying the nature of the human agent; humanity is a kind of *good* character. Virtue is therefore thematically pervasive in the way that we organize perspectives of the world through narrative. This is the enduring value that is *ritualized* in a genre that is uniquely preoccupied with performing what it means to be a human being in a world. To expand on Kupfer's argument: while narrative arts emphasize "the character of their characters," it is science fiction narratives that emphasize the human-ness of their humans. Narratives imply a sense of what our exemplars *should* look like even when our protagonists fail to embody them. There is a kind of tragic nobility in pushing back against an existence which challenges our moral development and resists our potential to fully actualize our humanity.

While this account of virtue offers advantages over the various forms of MPS, it is reasonable to be cautious of relying too exclusively upon the didactic power of narrative in order to ground the aesthetic value of our moral ideas. For one thing, we understand that the elements of narratives have not been selected and organized based on truth correspondence, but on their capacity to produce specific aesthetic effects in audiences. A narrative, even one which intends a historical kind of factuality, must introduce its facts from a perspective, and thereby exclude something of the world in its depiction. A successful narrative does not present itself as an exhaustive litany of such facts, but as a plot — a selected organization of story elements into a sequence which may exaggerate or edit events in the interest of telling the story more effectively. This is usually not reflective of a natural sequence, which is a problem for didacticism. If we see

moral value in a narrative because it allows us to rehearse moral dilemmas that bear some significant resemblance to those in the real world, then it is a complication if such a narrative were to falsify, distort, confuse, or outright misrepresent the facts of these cases. This critique follows the logic of Plato's mirror argument: mimesis is an inferior way of reproducing the world and so it is useless to look to art as a guide for our actions.¹⁹⁴

Recall that Nietzsche's response to this argument is fairly simple. The primary function of art is not epistemic, and yet art is the only means by which an otherwise meaningless existence is made enduring.¹⁹⁵ Any inaccuracies in an aesthetic representation of the world do not necessarily conflict with this aim. In fact, remember that truthfulness, in its plenitude, is harmful to us. However, an excessively delusional perspective of the world is not sustainable either. Art that succeeds as an existential outlook must balance the *Apollonian* and the *Dionysian* impulses — the comforting illusion of structure coming into harmony with its antithesis — the overwhelming reality of the "primordial unity," in which structure and individuation are revealed as illusory. The problem with the *Apollonian* forms of pure narrative — drama, and lyric poetry — is that they inhibit a vital component of the meaningful aesthetic experience, and in them the value of the sublime is diluted. In *Dionysian* art, we become lost in the "intoxicating" effects of the primordial unity, embodied in music and dance.¹⁹⁶ Nietzsche praises the dialectic reconciliation of the *Apollonian* and the *Dionysian* that takes place in the form of Attic tragedy. The best works of tragedy must satisfy our intuition that there is nothing inherently safe or familiar that is necessarily woven into the fabric of reality, but they

¹⁹⁴ *Republic* §X

¹⁹⁵ BT §5; WP §822

¹⁹⁶ BT §7

must also provide us with enough of structure to orient ourselves in relation to them and allow us to make sense of their patterns.

I have argued that contemporary science fiction narratives can adequately fulfill the role of moral thought experiments because of the content of their narratives, primarily those that explore the nature of human beings and their relationship to a world. If such narratives could carry the weight of the Nietzschean “extra-moral” ethic — a way of living valued primarily on aesthetic grounds — then there would have to be something in these narratives that also reflects a range of dialectic tension between *Apollonian* and *Dionysian* elements. There are many ways that this could work. As a general example, consider a narrative which makes moral perspectives explicit in its plot, while also focusing on the sublime features of human existence that cannot easily be reduced to conventional logic or morality. The resulting tension between the dual approach could produce a narrative space in which the tragic condition of the human being is emphasized in an aesthetic way. In contrast with the more optimistic view of classical Eudaimonism, in the contrast of rational beings with an irrational or indifferent world, there is an implicit emphasis on the theme of meeting and overcoming resistance — which for Nietzsche, is a constitutive aim of human agency, and as such is one of the only universally prescriptive values.¹⁹⁷

While themes of ontology and overcoming resistance are common in the mythical narratives of Attic tragedy, Nietzsche’s appraisal of these works extends beyond the content of the stories they tell to the form in which their narratives are presented. The

¹⁹⁷ Katsafanas (2013, 146)

first and most fundamental argument that Nietzsche introduces in *The Birth of Tragedy* is that the *Apollonian* and *Dionysian* impulses which together give rise to Attic tragedy each manifest in very different kinds of formal expression, the *Apollonian* in the art of the image-maker and the *Dionysian* in the “image-less” art of music.¹⁹⁸ This synthesis is most salient where the “dream-images” of *Apolline* lyrical scenes give way to the very “un-*Apolline*” choral passages called *dithyrambs*.¹⁹⁹ While Nietzsche does not articulate in detail what the dithyramb sounds like, he does describe its effects: an ecstatic departure of form, inspiring some pure, immanent feeling above and beyond the experiences we rationally process. In this state, we are overtaken by the sublime beauty of that which threatens the comfort of our *Apollonian* veil.²⁰⁰ Authentic aesthetic experience is something that requires giving up control to this feeling, to some degree. We know that the *Apollonian* is partially “dishonest” in its presentation — the dreamlike illusion of a world. A purely *Apollonian* narrative may still employ its fictions as a form of ritual or rehearsal — as a thought experiment — but we would once again find ourselves in a regress as to what qualifies as the criterion of value for our ethical conclusions.²⁰¹ Our affective response to the sublime involves some level of releasing control to our impulses. In order for some moral outlook or way of living to hold aesthetic merit, we must be moved by its beauty at a deeper level.

To be clear, the specific choral device of the dithyramb is not the only expression of *Dionysian* influence in Attic tragedy, but as I have stated, Nietzsche presents it as one

¹⁹⁸ BT §1

¹⁹⁹ BT §2

²⁰⁰ BT §5-6

²⁰¹ See discussion in Chapter 2, regarding the critique of conventional models of foundational ethics and metaphysics.

of the most salient examples. As I argue films to be narrative forms that are particularly well-suited to demonstrate the aesthetic value of moral perspectives, I raise the question of whether there is something in these works that functions analogously to the Attic dithyramb. It is one thing to argue that the content of specific science fiction film narratives resembles human life in such a way as to test moral ideas through counterfactual scenarios, and another to state that we value exemplary human narratives in primarily aesthetic ways. Film narratives communicate their meanings to us in complex ways, and on multiple sensory levels simultaneously, both explicitly and implicitly. Films employ a variety of interrelated techniques to this end, many of which manipulate our senses and emotions, and in so doing, influence the way that we feel about a subject. The most obvious analog to the dithyramb is a film's musical score and ambient sound design, which may be activating feelings and responses to the narrative content at a level below our awareness. However, this is not the only technique that film employs in this way; various kinds of visual imagery and iconography, the "beats" and rhythms of editing transitions, a film's use of light and color, spatial relationships between figures, how subjects are framed — all of these elements, among others, taken together may represent the "tragic chorus" of a film's stylistic language.

As I have argued and now hope to demonstrate, a study of science fiction framed in this context will add new dimensions to a reading of Nietzsche's critique of ethics, which will in turn help us to articulate details of how the concept of humanity is conventionally constructed, the normative functions of this narrative construction, and its usefulness to us under the conditions in which we exist. It is fairly straightforward to suggest that film has the potential to be incredibly poetic in its expression. I will make a

slightly stronger claim: science fiction films — in particular, those works that most effectively represent their ontological themes — function as *Attic myths* of the contemporary era.

3.4 *Two Frankensteins: A Framework for Understanding the Human Narrative*

Though the genre category would not come into usage until a century after its first publication, Mary Shelley's novel *Frankenstein; or, The Modern Prometheus* is often retroactively regarded as the first work of science fiction. Whether true or not, this claim implies that the novel either introduces new kinds of narrative themes or explores familiar themes in new ways. To some extent, both of these conditions are true. Conceived in a post-enlightenment world, *Frankenstein* offers a uniquely dark (dystopian) view of a world well-aware of new technological capacities to actualize many things that were previously impossible. The novel, like many subsequent works of science fiction, reflects the technological dynamism of its world, exploring various concerns over where the trajectory of progress may take us.

Recall that for Philip K. Dick the distinguishing characteristic of science fiction is the broad concept of the “novel idea” which impacts the world of the narrative and provokes change in the lives of its characters. The reason that this idea is so often a technological element is that through technological progress, novel concepts transition from the realm of the fantastical to become practical, or even inevitable. As technology has become more sophisticated, so has our ability to technologically replicate, and in some cases replace various human activities. In fact, the rate of technological advancement in society is increasing along an exponential curve rather than a linear

slope, leading Ray Kurzweil to voice the concern that we will soon cross a threshold at which the impact of technology on human life will have permanently altered what human life actually is (if we have not already).²⁰² While the debate about the advantages and risks of technological *transhumanism* is ongoing, a relatively neglected implication is that the theoretical possibility of technologically-created beings simulating humanity would render the “human” a trivial categorization. Fittingly, John Wilson refers to this as “The Frankenstein Problem”.²⁰³ While a difference of causal origins (e.g., being born rather than being constructed in a lab) would still distinguish two such beings, this does not seem like a significant enough means of differentiation, if the beings are otherwise physically and functionally identical.

Wilson and transhumanist theorists who followed him are primarily concerned with the progress of technology ultimately outpacing us, replacing us, or rendering us obsolete in its wake; I am more interested in the ontological questions that the Frankenstein Problem raises: what is a human being, and why is this question so difficult for us to answer? Is there a meaningful way of distinguishing the human from that which simulates it? At this point, I should re-introduce a distinction proposed earlier. Many works of science fiction engage with the ontological question through narratives which juxtapose the human with the non-human in order to illustrate some criterion of difference. Often this distinction is explored by introducing a kind of being that simulates

²⁰² Kurzweil (2016, 146). This concept articulates concern over the “technological singularity,” the threshold of progress at which artificial intelligence will become so sophisticated that it permanently surpasses ordinary human intelligence. Theorists speculate whether this represents an existential threat for forms of humanity that do not progress concurrent with the new “Super” AI.

²⁰³ Wilson (1964, 223)

the form of an authentic human but fails to do so by notably lacking some features that the narrative implies to be *essential* to a more “proper” form of humanity.

The novel *Frankenstein* depicts the events leading up to and following the creation, through some obscure scientific process, of a being which is never named, and often only referred to as “the creature,” or by the protagonist as “the wretch” or “the monster”. This creature is a being that is consistently described as horrifying in outward appearance, and yet recognizable as approximately physically human, and displaying at least some of the same capacities and dispositions of human beings. Based on its physical differences alone, this creature could never be mistaken as naturally-human, and so in some sense at least, the “lack” which differentiates it from authentic humanity is clear. However, is the lack represented in this narrative construction *essential* to the distinction? Are these ontological differences significant or relevant to our concept of the human? Actually, we will find that an answer to this question is contingent upon which *version* of the creature we use as our ontological model — which Frankenstein are we talking about?

The way that the creature is depicted has evolved across the various adaptations and incarnations of *Frankenstein*, resulting in differences in how the narrative reflects intuitions about the concept of the human being. The earliest and possibly the most significant of these changes was undertaken by Shelley herself in the 1831 edition of the novel. While Shelley notes in the preface to this edition that the majority of her revisions to the original 1818 version of the text were mostly minor stylistic revisions, Despina Kakoudaki points out that it is the preface itself that introduces a noteworthy and

compelling editorial change.²⁰⁴ In addition to a brief account of the genesis of Shelley's idea for the novel, the preface includes commentary clarifying Shelley's attitude toward one of its more ambivalent plot elements:

Frightful must it be; for supremely frightful would be the effect of any human endeavor to mock the stupendous mechanism of the Creator of the world. His success would terrify the artist; he would rush away from his odious handywork, horror-stricken. He would hope that, left to itself, the slight spark of life which he had communicated would fade; that this thing, which had received such imperfect animation, would subside into dead matter [...]²⁰⁵

Shelley's response here becomes prescriptive for us — this is how we are intended to interpret Frankenstein's cruel rejection of his creation. The rejection is significant, because rejection appears to be the primary, if not exclusive, inciting event of the compounding tragedies that unfold over the course of the novel. Upon realizing the successful animation of his creation, Frankenstein recoils from its appearance:

Oh! No mortal could support the horror of that countenance. A mummy again endued with animation could not be so hideous as that wretch. I had gazed on him while unfinished; he was ugly then; but when those muscles and joints were rendered capable of motion, it became a thing such as even Dante could not have conceived.²⁰⁶

It is notable that Frankenstein's horror is grounded in a visceral response to the creature's appearance, on aesthetic terms; the creature is "hideous," "wretched," and "ugly." While the De Lacey family's later rejection of the creature is the event which provokes his first violent eruption, it is the earlier rejection by his creator, and the further discovery of the circumstances surrounding the initial rejection, which exacerbates the creature's rage. Without Shelley's commentary, Frankenstein's actions seem cruel and

²⁰⁴ Kakoudaki (2018, 298)

²⁰⁵ Shelley (2014, viii-ix)

²⁰⁶ Shelley (2014, 36)

inexcusable — a shallow disregard for the feelings and welfare of another being that is obviously capable of experiencing emotion, and that is also obviously in a position of need. We may very easily empathize with the creature’s motivations, even if we cannot sympathize with his later actions. However, when we consider the comments in the 1831 preface, Frankenstein’s actions in rejecting the creature are preemptively framed as understandable, conscionable, even expected under the circumstances.

Kakoudaki rightly interprets our response to the moral implications of Frankenstein’s actions as contingent on whether we prioritize the political implications of the creature’s treatment, or the ontological basis of justification for this treatment.²⁰⁷ Shelley obviously prioritizes the latter, which entails the view that by virtue of what kind of being the creature is, Frankenstein has no obligations toward its well-being. In the first place, the creature should never have been created — it has no *right* to exist. Since its creation has already occurred, the next best thing is for the creature to die.²⁰⁸ In the protagonist’s panic, lacking the resolve to perform any act of mercy killing when he has the opportunity, the only remaining acceptable outcome is its rejection, to be left to the elements, left to perish of neglect. The only way that Shelley can justify this apparent cruelty is to regard the creature as fundamentally non-human, and therefore not deserving of the same social acceptance and humane treatment.²⁰⁹ The problem with this view is that despite its violent tendencies, the creature exemplifies many characteristics that we associate with a post-enlightenment human ontological narrative contemporary to Shelley. The creature is not just a mechanism; it has agency, and a sophisticated capacity

²⁰⁷ Kakoudaki (2018, 296)

²⁰⁸ The parallel to the myth of Silenus to this implied moral position is noteworthy.

²⁰⁹ Kakoudaki (2018, 296)

for both reason and emotion. It is not even clear that the creature's violent tendencies and impulsiveness are properties that should necessarily exclude it from the category of *human*, as these characteristics are also common in human beings. It is likely that Shelley was aware of the effect of these traits in her characterization but deemed them less important than the contribution of specific values to a presumptive ontological (and by extension, political) status for the creature. Shelley's ontological view is grounded in a commitment to a religious view, in which the very existence of the creature is an abomination, and so it is afforded no basic human rights.

Taken independently of the preface, the narrative does not automatically prescribe approval of Frankenstein's rejection of the creature. As Kakoudaki observes, the "social death" of the creature appears as an arbitrary punishment, and the ontological justifications of this punishment resemble many of the supposed 'arguments from nature' that ground racist ideologies."²¹⁰ Whether or not we were to interpret Frankenstein's initial act of creation as a transgression against nature, his subsequent actions would still be morally objectionable by this criterion.

The ambivalence of the Frankenstein narrative is reflected in a range of different characterizations and adaptations, varying in degrees between two distinct moral positions: the creature as an abomination or monster, and the creature as an authentic human subject. To make a conceptual distinction, I will refer to the narrative type which presents Frankenstein's creature as a pure "monster" with the phrase "*m*-type narrative" and the type which presents the creature as fundamentally human or *human-like* in some

²¹⁰ Kakoudaki (2018, 350)

significant way as an “*h*-type narrative”. Generally, science fiction narratives — those which take as their *novel* idea the synthetic being simulating the authentic — resemble some mode or mixture of the *m* and *h* narratives, and the primary ontological themes explored in them.

The influence of the *m*-type narrative on the science fiction genre results in works that emphasize the danger of ontological difference, including such films as: *Metropolis* (1927), *2001: A Space Odyssey* (1968), *Westworld* (1973), *Alien* (1979), *Saturn 3* (1980), *The Terminator* (1984), *Hardware* (1990), *The Matrix* (1999), among others. The effect of this emphasis is usually an overlap with the genre of horror. The prevalence and popularity of these films indicates our apprehension about a society in which artificial beings exist, and about the potential threat that they would represent to us. It is not just the case that these beings may not be capable of authentically instantiating the fundamental characteristics of a human being — they may not wish to. For such a being, there may be nothing of value in humanity.²¹¹

The most influential example of the *m*-type narrative may be the 1931 film adaptation of *Frankenstein*, directed by James Whale. While this film has many levels of aesthetic merit, it must be noted that in its adaptation, it has extracted much of the ontological and moral ambivalence present in Shelley’s novel and adopted the same sentiments toward the creature expressed in Shelley’s preface. In the film, the creature simulates humanity in form only, and only this to a certain degree. This narrative portrays the creature with an abnormality or lack of essentially human characteristics, thus

²¹¹ Once again, this is the concern addressed in debates about the technological singularity.

rendering it as a monster. The specific characteristics that the creature appears to lack are dispositions toward higher reasoning, self-restraint, and impulse control, and together these deficiencies frequently result in the violent and destructive expression of its rage. Because of this, the film is primarily an ontological narrative of difference rather than of same-ness. We never approach any true ambivalence about this creature's humanity; it is presented as more animal than human, owing to the flawed range of human capacities. In the film, these essential flaws are a symptom of a defect in physiology – specifically, the creature is endowed with a brain harvested from the body of a murderer. Because of this, the creature can never achieve the status of an authentic human subject. In fact, it may not even satisfy the conditions of higher life, above simple instinctive mechanism, as Frankenstein (here renamed as Henry) asserts: “That body is not dead. It has never lived. I created it.” The creature then both is and is not reducible to the aggregate of its cadaverous composition. It is both something new that transcends the dead pieces of other human beings of which it is constituted, and it is also something that is fundamentally *lesser* than the origin of any of those parts. Yet obviously these components have some effect on how the creature is endowed with different properties as well as limitations. The being itself is necessarily a monster — animalistic, impulsive, a *bad thing*. When Frankenstein asserts, “It is only a piece of dead tissue,” he is warned: “Only evil can come of it.”

Early in the film, as we are first introduced to the newly animated being, its face is illuminated primarily by directional lighting from above and below, both of which cast extended shadows that distort its features, emphasizing the iconic wide forehead and deep brow. The mechanical uniformity in the flattened dome of the creature's cranium and

protruding neck bolts stand in striking dissonance with the natural unevenness of its pale decaying flesh lined with sutures. The effect is of a kind of “steampunk” cyborg. The human being has been reduced to the animal, the animal ultimately reduced to the material, and the material intersecting with the technological. The film effectively represents a being far-removed from the domain of nature, and of life. It evokes a world composed of death — mechanical, decaying.

Nevertheless, the film does communicate a shifting level of sympathy toward its version of the creature in the way it is presented. When we are first introduced to the creature in the darkness of Frankenstein’s laboratory, and then in the dungeon-like cell in which it is confined, the effect of chiaroscuro lighting is to visually exaggerate its inhumanity. Later, when the creature emerges into the world, and we see it in daylight, the creature appears more naturally to us. With light filling in its features and giving its facial expressions a more “normal” human appearance, there is more reason to view the creature as endowed with emotion, and even a sense of wonder with its world. Having viewed the creature this way, its treatment at the hands of the human mobs evokes the revulsion we normally associate toward those who perform acts of cruelty. We sympathize with the creature’s suffering. The sympathy we may experience toward the creature does not likely extend to the point that we acknowledge the creature’s status as an authentic human being. Much like an animal, the creature’s violence is not premeditated, but reactive. Ultimately, the film’s narrative upholds the creature’s destruction as a tragic necessity, as the removal of a threat to the ordered and rational human world.

In the end, the moral focus of the narrative is on Frankenstein himself — a reflection of a human being absorbed by his own hubris nearly to the point of self-destruction. As the apotheosis of the “mad scientist” archetype, Frankenstein intends the monster’s existence to glorify his own, but in the process, he has created a being that only knows conditions of abuse and rejection, and often reacts to these conditions with violence toward human beings. There is a pervasive moral theme in this narrative, but it is not about the arbitrariness of social difference, or the belief that all sentient beings are afforded the same basic treatment. In other words, this is not a narrative of monsters revealed to be humans, but a warning against humans thoughtlessly *increasing* cruelty in the pursuit of egoistic aims.

3.5 Searching for “Human” in the “In-human”

While the 1931 film adaptation of *Frankenstein* effectively demonstrates the common patterns and themes of an *m*-type narrative, its source novel *Frankenstein* is in many more ways the prototypical *h* narrative type. The novel reflects a more sympathetic view of the artificial subject, as unlike the film’s version, this creature is a being with the legitimate capacities for complex emotions, advanced reason, and drives toward flourishing and happiness. In the novel, details of the creature’s creation are no less ghoulish, intended to frighten or disturb readers, but without the frame of Shelley’s preface they do not necessarily stand as the affront to nature that Shelley presumes. We cannot help but observe that the conditions of the creature’s emergence and developing consciousness bear significant similarities to our own. When human beings mature to a certain level of awareness, we discover that we inhabit a world not by choice, but by circumstance, and that we are subject to factors of suffering, neglect, and loss which are

largely beyond our powers to control. In spite of these conditions, Frankenstein's creature frequently exhibits the disposition to affirm life as inherently meaningful, even in the face of ubiquitous suffering and loneliness. As an example: upon his tragic rejection by the De Lacey family, the creature briefly contemplates suicide, but is so affected by the inherent beauty of nature — of the world itself — that he withdraws from the abyss of nihilism.²¹² This pattern recurs, with the creature experiencing isolation, fear, and despondency, only to afterward be moved by a feeling of the sublime value of existence. When confronting his creator, the creature explains, "Life, although it may only be an accumulation of anguish, is dear to me, and I will defend it."²¹³ The existential element adds nuance to the creature's characterization and is deeply ingrained in the narrative fabric of the novel.

While the creature perpetrates horrific acts, the acts are motivated by impulses that are humanized, by way of the precipitating conditions of provocation and abuse. For these reasons, among others, we can conclude that the narrative accepts some degree of the proposition that authentic humanity results from properties of character rather than circumstances of one's biological origins, and that the former is not dependent on the latter. Because the novel *Frankenstein* seems to indicate that a range of moral character is possible in both human and non-human agents, this opens up the possibility of some crossover of these properties, which frustrates attempts to *essentialize* any of them in the conception of the human narrative. When we consider one of the fundamental suppositions of the *m*-type narrative — that artificial beings of science fiction are usually identified by some *difference in* or *lack of* characteristics that distinguish them from the

²¹² Shelley (2012, §16)

²¹³ *Ibid.* §10

authentic humans of the narrative — the alternate *h*-type narrative often implies any such supposed difference to be arbitrary. An *m*-type narrative may uphold differences with the intent of justifying the harm or exploitation of non-humans in the narrative — a rhetorical tactic which often reduces the chance that we will empathize with the *other*, the monster. But an *h*-type narrative acknowledges the intuition that our empathy toward the other may be well-placed. In genre terms, the characteristics which we presume to be conditions of authentic humanity may exist in a subject that is not actually biologically human. Similarly, our intuitions may be confirmed when a subject that is biologically human may come to lack these presumed characteristics, demonstrating a failure to maintain their “human-ness” in some meaningful ways — a *moral* failure. By prompting an emotional response to the treatment of artificial subjects, these narratives communicate a moral perspective on the question of what it means to be human.

Artificial humans present counterexamples to the standing arbitrary criteria of *unlikeness* — they function as a kind of narrative proxy for oppressed groups, a critique of how human beings ground structures of oppression in arbitrary criteria of difference. Ridley Scott’s *Blade Runner* (1982), loosely adapted from the Philip K. Dick Novel *Do Androids Dream of Electric Sheep?* is one such narrative.

Blade Runner is a film which employs the *h*-type narrative paradigm to challenge a conventional ontological model: that of a human being as necessarily a naturally occurring biological organism with special properties of character that are unique to this state. By juxtaposing human characters with artificially created *human-like* androids called “replicants,” the film criticizes the basis on which we would view these beings as necessarily inferior to their human counterparts. Like the creature in Shelley’s novel, it

becomes apparent that it is at least possible for replicants to possess or develop those unique traits that we associate with *human-ness* in the most important sense, whether they qualify as biologically authentic or not.

The replicants of *Blade Runner*, like the synthetic beings of many works of science fiction, face negative treatment because of their status as non-human. In *m*-type narratives, this treatment is often justified on the grounds that “counterfeit” beings are deservedly designated as the antagonists of the narrative — they are the aforementioned monsters, villains, or are simply mindless mechanisms performing some default routine which is harmful. Harm may not be the primary motivation of such villain types, but a secondary consequence of some primary aim. Even the purely destructive T-800 “cyborg” in the film *The Terminator* is under the control of a utilitarian intelligence which views the destruction of humanity as fundamentally good above and beyond the harm it inflicts on those humans. Films like this are simply a high-tech version of the *m*-type Frankenstein, embodying some cultural anxiety we feel over whether futuristic “super-AI” would have any logical reason to uphold values which are compatible with the good of human beings. By contrast, in the case of *Blade Runner*, the replicant antagonists are motivated by primary goals similar to our own: their intention is to survive, and to enjoy some level of a flourishing life. Sometimes, much like humans, replicants may inflict harm in pursuit of their primary aims, but this does not distinguish them as monsters.

Because of their degree of similarity, grounding a distinction between authentic human beings and replicants becomes necessary to a world which would maintain the mistreatment and lower status of the latter by the former. Humans in *Blade Runner* use a

justification of ontological difference, with the intent to reframe what we would otherwise recognize as violence, exploitation, and abuse as an acceptable matter of course, by virtue of what kinds of beings replicants are. If replicants are objects in the purest sense — mere tools of humanity — then there is no abuse, as there is no actual subject to experience this abuse. It would be strange to bestow human rights on a very sophisticated refrigerator, or a smart-phone, for instance.²¹⁴ However, as Alf Seegert notes, if replicants and humans share features that imply a fundamental commonality, then it becomes more difficult for humans to ethically justify their mistreatment, and their relegation into a lower class of being.²¹⁵ To add to this, even if replicants are somehow “less” human in some way, but still capable of experiencing suffering, it would be easy to argue that it is unethical to *cause* them to suffer.

In many ways, replicants do not appear to be as dissimilar to human beings as the conventional distinction would require. In fact, their function in service to human beings relies upon their similarity in some significant ways. These similarities are likely engineered to facilitate the various functions for which replicants were created, all related to their service as a proxy worker class. Replicants fulfill a variety of specialized servant roles, including those explicitly mentioned in the film: soldiers, manual laborers, and sex workers. Marilyn Gwaltney points out that in order to perform these roles, it is likely that replicants require both a physical form similar to humans, in order to carry out the same

²¹⁴ I should acknowledge that Kurzweil’s theory holds that my comment here may not age very well.

²¹⁵ Seegert (2011, 40)

kinds of tasks, and some degree of autonomy and the capacity for advanced reason, to effectively negotiate challenges related to these tasks.²¹⁶

The existence of replicants reflects a feature of a very specific reality for humans in the setting of *Blade Runner*'s dystopian Los Angeles of 2019. In this world, humans have damaged the world to the point that the survival of a functional society is threatened.²¹⁷ The environment is toxic — a detail underscored by the complete absence of biological animal life, and the markedly *degenerate* state of the human life that remains. If biological life can no longer thrive in this environment, then the artificial life that supplements its labor would need to be much more resilient.²¹⁸ Because of this, the most noticeable physical differences between the two types of beings appear related to the specific ways in which replicants *excel* at humanness — their capacity to be “more human than human,” a question-begging pronouncement of their ontological condition by the Tyrell Corporation that produces them. One of the ways that this is most evident is in the physical strength of replicants. It is also established that replicants are manufactured to perform those tasks associated with life on off-world colonies which are carried out in conditions adverse or unfavorable to biological human beings. This means that replicants are designed to endure conditions of suffering. But how do we know that they actually suffer? How do we know whether they are nothing more than mechanisms sophisticated enough to perform their tasks effectively, or whether their resemblance to humans also extends to the capacity for authentic consciousness and feeling?

²¹⁶ Gwaltney (1991, 35)

²¹⁷ *Ibid.* 32

²¹⁸ *Ibid.* 35

Our best evidence that replicants are capable of suffering is their responses to conditions of suffering. The physical resilience of replicants does not appear to mitigate their suffering in any important ways. We see the replicant Pris bare-handedly pull an egg from boiling water without suffering any apparent injury or pain. We also see her affected by the same existential distress of loss and finitude which plagues human existence. We see Pris and her companion Roy Batty express resentfulness at the humans who uphold their subjugation and impose their suffering. The narrative portrays these replicants in ways that imply their actions to be motivated by real pain, rather than simulated. Like us, they not only have the capacity to experience suffering, but the sophisticated powers of reason through which they may also question the circumstances of their suffering, as well as a variety of other factors related to the conditions of their existence. In *Blade Runner*, replicant characters are motivated by a will to survive and to avoid mistreatment, to flourish, and they appear to be endowed with the characteristically “human” capacity of reason with which to pursue these ends. The problem is that for replicants these pursuits are constrained by their social role, which involves *social death* — the complete lack of any intrinsic rights that personhood entails. To justify the exclusion of replicants from human rights on the basis of being *different* from the category of humanity would require a convincing account of what this difference consists in.

To explain how *Blade Runner*'s ontological themes function as a kind of *reductio ad absurdum* a conventional human narrative, Ross Barham discusses the supposed differences between humans and replicants in terms of Aristotle's *Four Causes*. The differences that the efficient, the material, and the formal causes reveal are essentially

incontrovertible, on the basis of the artificial origins of replicants; replicants are designed in the form of humans and produced by the Tyrell corporation out of various synthetic materials, while biological humans are born and develop through natural biological processes, and are composed of organic matter such as flesh, cells, and proteins.²¹⁹ In order to establish that there is no meaningful ontological difference between the two categories of beings, we would have to hold these causal differences to be insignificant. This is simple, because the form of the argument already presumes that these differences are a condition of the comparison. The premise that I begin from in this chapter can be stated in the following way: the narrative of the film *Blade Runner* persuades us that artificially created human-like beings are not necessarily distinct enough from true biological human beings to warrant their mistreatment on the basis of this distinction. While a variety of theories could discount the significance of the efficient, formal, and material causal differences between replicants and humans, it is beyond the scope of the current argument to explore the merits and problems of such theories. Instead, I am concerned with the implications of applying Aristotle's *final cause* — the purpose of a thing's existence.

Aristotle's final cause, or as it is applied to his study of human ethics, the *function argument*, is significant, because when we consider the narrative of *Blade Runner*, the argument highlights the most prominent differences between human beings and replicants.²²⁰ Because they are created for the purpose of performing specific service tasks, it is reasonable to conclude that the *final cause* of replicants is to provide labor —

²¹⁹ Barham (2011, 15-16); Aristotle, *Metaphysics* V.2

²²⁰ Aristotle, NE §X

to serve the will of human beings.²²¹ Because there are no such superseding duties or demands universally placed on the activity of human beings, the function of a human being is less obvious to identify, but Aristotle has an answer. By virtue of being human, we are endowed with kinds of rational capacities, and our function — our *final cause* — is to develop these capacities, and through cultivating the right dispositions, express them in rational activity.²²² According to Aristotle, our function is to be happy, to flourish, to enjoy the most sustainable state of *eudaimonia* possible, and when we are engaged in rational activity is when we are most consistently in this state (footnote: do not worry for the moment that Nietzsche convincingly argues this aim is not a sustainable goal.²²³ The problem should be obvious now: replicants regularly demonstrate those same rational capacities, and in some cases, the same cultivated dispositions toward those activities and states that supposedly denote well-functioning humanity.²²⁴ The fact that replicants were intended to be exploited makes no difference if they are capable of the same suffering and striving that we are, and it is just as reasonable to presume that their ideal function consists in the same kinds of activities as ours does. Aristotle contradicts his own view, by dismissing the capacities of different groups of human beings that no doubt possess those capacities that he denies them. Like many of the characters in *Blade Runner*, Aristotle takes the untenable position of justifying the subjugation of groups of persons on the grounds of denying their full humanity.²²⁵

²²¹ Barham (2011, 15)

²²² Aristotle, NE §X

²²³ Do not worry for the moment that Nietzsche convincingly argues in *The Birth of Tragedy* that this aim is not a sustainable goal.

²²⁴ Barham (2011, 16-17)

²²⁵ Aristotle, *Politics* 1254b

A good argument can be made that the shared capacity for higher rational functioning establishes a fundamental similarity between humans and replicants, and following from Aristotle's function argument, that there is no good justifiable reason to exclude beings like replicants from participation in the same rights afforded biological humans. However, the humans in *Blade Runner* who hold replicants to be a *lower* class of being do not dismiss the fact that replicants are intelligent, or even autonomously rational. Like many other science fiction narratives, in *Blade Runner* it is fairly obvious, even expected, that artificial human beings are sophisticated AI that can function at a high cognitive level at least equal to that of human beings. Instead, the distinguishing characteristic which replicants are believed to lack, and which designates them as *lesser* beings on that basis, is the capacity for empathy. In theory, empathy as a necessary and sufficient condition of authentic humanity serves the narrative model well, as it requires a sophisticated level of cognitive and emotional functioning and facilitates the development of moral reasoning.

Empathy is a specific kind of reaction to our understanding of the experiences of another. According to Gavin Fairbairn, it is a feeling related to, but distinct from, the feeling of sympathy. Whereas sympathy is a kind of affective response to the experiences we perceive that another is undergoing, empathy involves a reaction at a more cognitive level, in which one can imaginatively project oneself into their situation, to grasp the significance of their circumstances, and in this way gain some genuine understanding of how they are affected by them.²²⁶ A replicant displaying agency and a full range of rational capacities, but lacking the capacity for empathy, seems conspicuously non-

²²⁶ Fairbairn (2009, 192)

human because of this lack, making it easier to justify them as a threat — as potential monsters. If replicants cannot genuinely empathize with others, then how can they really act in the best interests of those others? Following from this scenario, then those replicants that cannot be compelled to conform to their social role, or in this case, those that have rebelled from this role, necessarily represent a threat to human society if for no other reason than the fact that without the capacity for empathy, they could not even accurately process what constitutes harm toward others in order to reliably avoid it.

Blade Runner lures us into an *m*-type narrative with a familiar setup — the “technological horror” trope of androids as unfeeling machines. Over the course of the narrative, the film subverts many of the conventional expectations of this narrative type in order to establish itself as an *h*-type narrative. At the beginning of the film, a small group of rogue replicants from “off-world” have escaped and returned to Earth. The film’s opening crawl informs us that this is a common occurrence in this world, necessitating a specialized kind of police force known as a “blade runners,” who are tasked with hunting down and “retiring” these fugitive replicants. Because replicants are otherwise indistinguishable from human beings in most important ways, the blade runner’s primary tool in identifying them is an apparatus known as the “Voight-Kampff test” (VK), meant to determine whether a subject has the capacity for empathy. The test achieves this result by measuring involuntary physiological responses to questions designed to provoke feelings of empathy in the subject. This method is presumed to be effective, because no replicant in the film is capable of passing this test.

At the beginning of the film, a blade runner named Holden is applying the VK test to a subject named Leon — the first of each that we are introduced to: blade runner and

replicant. Leon is boorish and argumentative in the interview part of the test, but not outwardly threatening, at least at first. Both characters are framed in medium closeup, and the scene unfolds with a conventional pattern of crosscutting, back and forth between the two. Eventually, Leon becomes nervous, and questions the validity of the interrogation. As Holden's questions become more probing, we see Leon's dismissiveness give way to confusion and agitation, and then, something changes in Leon. Framed in closeup, Leon's features contort in distress as he struggles with one of Holden's questions.²²⁷ He flinches. Holden senses this tension in his subject and leans back into his chair slightly, drawing from his cigarette. "They're just questions, Leon," he maintains, disarmingly. As Holden casually proceeds with his questioning, our perspective cuts to a medium long shot perpendicular to the axis of action — we see both characters framed on opposite sides of the conference table that separates them, emphasizing a distance between them that was not obvious in the earlier shots. Holden asks a seemingly innocent question about Leon's mother. Leon responds by abruptly firing a pistol he has concealed below the table, striking Holden in the chest and spinning him around. The shot sends Holden's body, chair and all, crashing through the wall behind him. In a quick cut, we see Leon standing now, pistol pointed directly at Holden's back. He fires a second time, knocking Holden against a desk in the adjoining office.

The abruptness of the way this attack is presented results in a tonal clash with the deliberate and almost mundane pacing of the exchange which precedes it. The transition is intentionally disruptive, irrational — the shift is confusing to us. It is possible, as John

²²⁷ In the specific line of questioning, Holden is asking Leon to imagine a tortoise on its back in the sun, and to imagine himself not helping it.

David Ebert suggests in his analysis of this scene, that Leon reacts as he does because Holden has offended him with this particular question; strictly speaking, as a replicant, Leon has no mother.²²⁸ However, there is a more likely interpretation. Leon does not attempt to kill Holden in retaliation for his offensive line of questioning, or even out of casual machinelike carelessness. His actions are the result of a failure to process that which exceeds the limits of his functioning — empathy. At this point in the narrative, we see Leon’s lack of empathy as a legitimate danger, and the scenario of replicants hiding among us as a version of the “wolf in sheep’s clothing” narrative trope. But it is not Leon’s inhumanity which makes him dangerous. He would have likely acted a different way if he were not backed into a corner like this, and if he were not well-aware that his test results were betraying him. Leon bears more resemblance to the desperate “outlaw fugitive” trope than the “cold cyborg menace” trope.

If we accept the premises that only humans can truly exhibit empathy, and that the VK test accurately measures empathy, then it follows that replicants are not human. However, the argument fails on the basis of both premises. Over the course of *Blade Runner*, the empathic capacity of replicants is depicted as plausible, no less so than that of human beings.²²⁹ The VK test has no real way of accounting for this. Replicants react to the deaths of other replicants with what appears to be authentic remorse, and in one notable instance, the replicant Roy Batty saves the life of the protagonist, the very blade runner who has pursued him, and who is also directly or indirectly responsible for the death of every one of his companions. These facts challenge the reliability of the VK test

²²⁸ Ebert (2015, 29)

²²⁹ Seegert (2011, 40)

to determine an authentic capacity for empathy. While replicants are incapable of passing this test, ostensibly indicating a deficiency or lack of this essential human characteristic, through their actions, at least some of the replicants demonstrate examples of empathy in practice. This implies that the secondary physiological response to situations that provoke empathy, and the actual feeling and expression of empathy itself, may not be reducible to the same thing. If there is no necessary connection between the physiological response that the test registers and the actual emotion, then which one is the test measuring? Perhaps it is measuring something distinct in the different ways that biological and artificial beings process stimuli that trigger empathy. If the test works this way, which appears to be the case from what we see in the film, then it is establishing a marker of physiological difference, without accounting for the fact that this difference is a functional irrelevancy. Either way, this would be a poor method of quantifying the authenticity of human beings. If *true* empathy, like agency and reason, is realizable in artificial humans, as the narrative implies, then it also fails as the criteria of distinguishing between humans and replicants. If the difference is actually only the physiological difference in the way that the same feeling is processed by two different kinds of systems — one of biological origin in the first case and bio-mechanical origin in the other — then the test is likewise irrelevant.

To be predisposed toward cruelty and to lack the capacity for empathy are not precisely the same thing, but we usually view an empathic disposition as something that precludes the tendency toward cruelty, because of how empathy gives us a more complex consideration of the impact of cruel actions on others. In other words, if we know enough about the experiences of another, then we would not wish the other to suffer. This may

not hold as true out of necessity, but we conventionally take empathy to be the kind of trait that leads us to care for the best interests of others, to not wish them to suffer, and not to knowingly cause their suffering. Human structures of authority in *Blade Runner*, represented by the characters of the three blade runners and their overseer, Captain Bryant, systematically cause harm to beings who are evidently capable of experiencing suffering, with the VK test as the sole justification that these instances do not actually constitute cruelty. The plight of the escapees gives us a broader sense of the cruelty inflicted by human society at large, and the absence of what we view as an intuitive level of empathy.

The object of our empathy shifts over the course of the narrative. While we do not know very much about Holden, he is characterized as more of a bureaucrat than an assassin, and in this context, Leon's actions appear brutal, and difficult to justify. Later, when we consider this scene within the context of the plot, we can see that Leon's interrogation is nothing less than a defense of his life, and his actions are a last resort when this defense fails. For blade runners, there is a casual, administrative detachment to the way that they approach the retirement of replicants that is intended to unsettle us. It is possible that the replicant Leon is not actually capable of empathy, but this does not single him out as necessarily sub-human on this basis alone.

As a protagonist, we expect that Holden's successor Deckard will rise above his reluctance of duty and human vices to ultimately comport himself heroically, morally. However, we quickly see that there is nothing heroic in the way that Deckard carries out his duties. While a resourceful detective, Deckard is depicted as mostly inept as an assassin, only able to subdue his targets through luck or the intervention of others. When

he does finally manage to “retire” the replicant Zhora, the scene plays out in a way which emphasizes her suffering and suggests Deckard’s cowardice and cruelty. As Deckard fires on Zhora, striking her repeatedly in the back, Zhora crashes through one glass door after another in slow-motion. We are not accustomed to seeing the *human* way in which the android antagonist suffers and dies, or the way that the human protagonist is *de-humanized* in his careless disregard for her fear and pain.

As Stephen Mulhall argues, scenes such as Zhora’s death, as well as other scenes depicting violence inflicted upon replicants, trigger the same kind of affective response as violence inflicted upon human beings.²³⁰ This response does not rely upon confirmation of authenticity, of the being inhabiting this human-like form — it is “behavior of a particular complexity” which settles this matter for us prior to questions about deeper levels of difference.²³¹ The form of replicants is human enough to trigger our empathy when they are harmed and killed; they “read” as human to us before we can attempt to rationalize whether they are or not. Their effective powerlessness to resist makes the violence more impactful, allowing us to more easily recognize that the replicants are victims, and not aggressors.

All of this suggests that an intersection of types is taking place — replicants are beginning to demonstrate the capacity for empathy, while humans are beginning to demonstrate a lack of this capacity. The cruelty of humans is emphasized, while replicants are depicted as violent only out of desperation. Even Roy Batty’s “cat and mouse” sequence with Deckard at the end of the film resolves in one of the most

²³⁰ Mulhall (1994, 88)

²³¹ *Ibid.* 88-89

fundamentally *human* moments for both characters. As Deckard flees Batty, he leaps from one rooftop to the next, but misses his landing, and ends up clinging to an outcropping of the rooftop. Deckard struggles to pull himself up, but is obviously losing this struggle, inching gradually closer to his death. Batty watches Deckard struggle, letting him dangle a moment longer. He smiles as Deckard's fingers slip. Only when Deckard's grip fails completely does Batty snatch hold of his wrist, effortlessly hauling him up one-handed and letting him hang there a moment. He tosses Deckard abruptly onto the roof and slowly approaches him. Deckard crawls backwards, fearfully, until he finds himself backed against the corner of a wall. He cowers as Batty crouches in front of him. The two sit facing each other through the downpour of the rain. Batty cradles a dove in his hands. When Batty speaks, Deckard regards him with a mix of confusion and terror: "I've seen things you people wouldn't believe," Batty begins. "Attack ships on fire off the shoulder of Orion. I watched C-Beams glitter in the dark near the Tannhauser Gate. All of those moments will be lost in time, like tears in rain."

We do not have a sense of the full meaning of the phenomena that Batty describes in his final moments before he expires of a premature old age, but that is the point; we have no point of reference for Batty's experiences. Everything he has seen represents something unique, something that he recognizes as beautiful, and as tragically bound in the temporality of his existence. His valuations of his life are on aesthetic terms, and they embody the principle of *amor fati* – the embrace of the circumstances of one's existence as fundamentally beautiful exactly as they are.²³²

²³² GS §278

In contrast to the arbitrary responses measured by the Voight-Kampff test, Batty's rescue of Deckard is established as the truest expression of authentic empathy in the film. Batty intervenes with Deckard's fall, in both a literal and metaphorical sense. Not only has Batty rescued Deckard's life, but his humanity as well. The impending ontological intersection, at which biological humans are becoming more mechanistic and machines are becoming more humanlike, is occurring at this moment. As Batty grows to embody authentic empathy, Deckard has grown alienated, disaffected, and callous in his regard for others – that is, until Batty rescues him from his descent and guides him through the ordeal which leads him to an understand of true empathy.²³³

This distinctive overlap is common in *h*-type science fiction narratives, in which artificially engineered humans (the supposed “counterfeit monsters”) may come to embody *more* of the distinguishing characteristics of humanity than their biologically human counterparts. In his essays *Man, Android and Machine* and *The Android and the Human*, Philip K. Dick expresses a growing concern over this inevitable point of convergence between the growing sophistication of machines and the gradual effective “de-evolution” of human beings in response to an alienating world, among other factors.²³⁴ In *Blade Runner*, the setting is unquestionably an alienating place — toxic, bleak, and impersonal. Human beings are becoming more disaffected, and ironically less demonstrably empathic, in a world which holds empathy to be the primary designator of real humanity.²³⁵ Meanwhile, the increasingly sophisticated replicants, originally

²³³ It makes no difference whether we accept the implied ending of the “Final Cut” version of the film – that like Rachel, Deckard is actually a replicant himself without realizing it. The primary reason that this does not matter is that Deckard has not suffered the same treatment for his status that Roy Batty has, having essentially lived his whole life in a “privileged” class.

²³⁴ PKD (1995, 183; 211)

²³⁵ Davies (2015, 140)

conceived as human-surrogate labor machines, are becoming rebellious, difficult to control, and agitated with their social designation.

If the same distinguishing characteristic (empathy) is present in some human beings and some replicants, but not universally in either, what does this mean for *Blade Runner*'s specific narrative of human authenticity? Having convincingly challenged the validity of the Voight-Kampff test, should we assume that the film also dismisses the criterion of empathy as a necessary or sufficient condition of humanity? This is not the case; instead, it is the distinction between human and *not-human* that is challenged. Berys Gaut proposes that while the film places the suppositions of two different ontological narratives in confrontation with one another, the distinction that this confrontation presumes is not actually between humans and non-humans, but between two different categories of the concept of *humanity*. Gaut refers to the first category as "biological humanity," a category which presumes that authentic humanity is determined by facts about the organic origin and composition of the human body. The second category, "evaluative humanity," refers to a set of character dispositions which human beings possess – specifically, in the case of this narrative, those toward kindness, forgiveness, and especially empathy.²³⁶ Because *Blade Runner* sets up empathy as the fundamental signifier of authentic humanity, it also presumes evaluative humanity and not biological humanity to represent the more authentic version of the category of human being. Evaluative humanity is a kind of narrative by which we account for the "virtue set" of the human exemplar. Through this category, we also define the nature of an authentic human being not by arbitrary material composition, but by patterns of action which constitute

²³⁶ Gaut (2015, 35)

kinds of character. If, as argued earlier, agency and action are functionally inseparable, then it makes sense to prioritize the evaluative criteria over the biological.

Through the acquisition of various dispositions of character, individuals who are born biologically human may develop and refine the types of virtues that lead to evaluative humanity. *Blade Runner* reveals to us that participation in the first category of humanity does not automatically entail participation in the other. It is possible for biological humans to fail to instantiate the category of *human-ness* in the evaluative sense — either failing to develop these traits of character (virtues) or by “losing” them in some way. In other words, beings that are born into the category of biological humanity might never achieve full evaluative humanity. Further, beings that are categorically *not* biologically human but possess the right kinds of basic capacities, such as certain replicants in the film do, may come to develop evaluative humanity. The specific range of human qualities – the aforementioned virtue set which constitutes evaluative humanity (again, kindness, forgiveness, and empathy) – is therefore a project of cultivating the right dispositions, similar to the process by which Aristotle accounts for the development of the virtues.²³⁷

3.6 Color and Authenticity in Ex Machina

Like *Blade Runner*, the plot of Alex Garland’s *Ex Machina* focuses on technologically simulated humanity, and explores philosophical issues related to a narrative of personhood and the values that follow from this. In the plot of *Ex Machina*, a young computer programmer named Caleb is invited to the isolated compound of a

²³⁷ Aristotle, NE 1103a30

search engine mogul identified only as Nathan, supposedly for the purpose of administering a Turing Test. The subsequent test consists of six interviews with Ava, a female android that Nathan has created in secret. As the series of interviews progresses, the relationship between Caleb and Nathan deteriorates, largely due to Caleb's growing mistrust of Nathan's motives, as well as his resentment over Nathan's mistreatment and objectification of Ava. *Ex Machina's* plot centers around a familiar genre convention: Nathan's shortsighted ambition leads to his ruin; he is the equivalent Victor Frankenstein of this narrative. The success of Nathan's innovation is contingent upon Ava being an authentic artificial intelligence — for all purposes, a person — and yet Nathan consistently treats her as a servant or a tool, and as fodder for his amusement. This either indicates that Nathan ultimately rejects his own assertion of having achieved “Strong AI” in Ava's creation, or his actions are indicative of how he treats actual human beings.

Ex Machina employs a range of specific visual techniques to manipulate our sympathies and establish character motifs. One of the most noticeable strategies by which the film influences our acceptance of Ava and other characters as authentic persons is explicit in the film's plot: color. Early in the film, yellow is the color motif most closely associated with the character Caleb. Soon after their first meeting, Nathan introduces Caleb to his sleeping quarters within the compound — a windowless room dominated by a saturated yellow tone. As Caleb reads aloud the legalese of the non-disclosure agreement he is being pressured to sign, Nathan reclines inattentively across a bed covered in a red bedspread – the most notable color contrast in the scene. As Nathan stands and approaches Caleb, his dark attire and beard cross into collision with the hue of the background plane, making him appear out of place with the environment. Caleb, on

the other hand, nearly disappears into the mise-en-scene, his blond hair corresponding to the yellow surfaces of the walls, the desk, the lamp, and the frame of a painting immediately behind him. In the shot which immediately follows this scene, the camera tracks across a wall completely covered in “post-it” notes. Almost all of the post-it notes are yellow, with a few random patches of red and blue. The camera continues to track until Nathan comes into our field of vision, positioned in front of an array of computer screens, his back to us. On the screen, he watches as Caleb enters the testing chamber, which transitions us to the first of the Turing Test sequences.

As the first Turing Test session begins, Caleb stands enclosed in a protective glass chamber within an underexposed room. Yellow highlights are reflected in the glass surface of Caleb’s chamber. Caleb’s attention fixates on a crack in the glass surface in front of him and he approaches it. A *rack focus* shifts our perspective to the foreground plane, revealing the contours of spiderweb cracks, illuminated in the center by a blue highlight reflected from a wall panel. As the camera slowly tracks around the chamber, Caleb runs his fingers deliberately across the blue reflected in the cracks. As he backs away from the wall, our perspective cuts to a window on the opposite side of the room from the glass chamber. Caleb stands at the right side of the frame, with his back to us. A silhouetted figure in the background enters the frame from the right, moving through the empty space to Caleb’s left, and stopping, framed in the blueish hues filtering in from outside of the window. As we cut to a reverse angle shot, we see Caleb facing the figure, and we see through an improbable gap in the figure’s torso. Instead of flesh, the figure’s midsection appears to be composed of metallic mesh and blue light. Caleb’s face registers surprise at the appearance of the figure, Ava. Blue lights on an electronic wall panel

occasionally flicker behind him, matching the constellation of blue lights within Ava's body, further emphasizing her physical artificiality. A vertical row of yellow lights ascends the anterior side of her spinal column — highlighting anatomy that would not be visible to us in a biological human.

In a later interview session, Ava reveals a drawing she has made. As she extends her arms to show it to Nathan, yellow tubing within her arms is prominently visible. As our perspective tracks around her, we see the yellow lights within her mesh body casing, and the chrome of her head reflects the yellow highlights of the testing room. In this particular interrogation, Caleb's wardrobe has shifted; instead of the muted gray, white, and black wardrobe of the earlier sessions, he now wears a blue shirt with a checkered pattern. Ava and Caleb's associative color motifs have become intertwined, as a visual representation of their growing fascination with one another.

In the fourth interview sequence, color as a character motif becomes more overt. During this session, Caleb introduces Ava to philosopher Frank Jackson's thought experiment of "Mary in the Black and White Room," also known as "The Knowledge Argument." As a challenge to hardline physicalist theories of mind — those which claim that all mental states are reducible to physical processes — Jackson proposes that a hypothetical subject (Mary) could be completely knowledgeable about every conceivable physical fact associated with the concept of color but having lived her whole life in a black and white room and never having experienced any color for herself, there would be something above and beyond these facts that Mary lacked. What Mary lacks is the

“qualia,” or *what-it’s-like* of the experience of color.²³⁸ In other words, conscious first-hand experience entails something more than simply being able to provide an account of that experience. Caleb explains to Ava that Jackson’s thought experiment signifies an ontological difference between a computer and an authentic mind. Whereas a computer can only process bare facts, human experience entails the component of qualia.

As Caleb relates this thought experiment to Ava, a slow zoom on a closeup of Ava’s face is crosscut with shots of Ava as Mary, confined in a purely black-and-white room. As this sequence continues, Ava-as-Mary emerges into the sunlight, seeing the blue and green hues of the landscape outside of Nathan’s compound for the first time. Ava’s parallel to Mary’s situation is obvious, but the crosscutting prompts a question: do we interpret this to be Ava’s fantasy? The slow zoom that continues across the cutaways gives us a sense that we are moving *into* Ava’s mind, or that Caleb’s words are prompting some inner realization in Ava. This “inwardly moving” perspective supports the conclusion that the cutaway shots are part of Ava’s internal diegesis, as she imagines herself in the role of Mary, implying an inner consciousness of some complexity, with the capacity for imagination.

The Knowledge Argument expresses skepticism over the proposition that an AI, no matter how sophisticated, could ever achieve something functionally equivalent to an authentic human mind. While the plot of *Ex Machina*, much like that of *Blade Runner*, challenges this conclusion, the fact that the argument uses the example of color is

²³⁸ Jackson (2008, 658)

significant. Color in *Ex Machina* has an associative effect — we come to relate specific characters to the color they are most often visually associated with.

In the film, Caleb and Ava have been associated with yellow and blue respectively, while Nathan is consistently associated with black, grays, and whites. At one point in their interview sessions, Ava conversationally accuses Caleb of lying when he says his favorite color is red. Periodically during these scenes, Ava secretly shuts down the facility's power, so that the two of them can converse free of Nathan's surveillance. These "power cuts" trigger backup lighting, casting the room in deep red. As Ava's question has hinted, red signifies deception — the two of them plot to deceive Nathan and escape the compound together during these "red-outs", but we also find out that Ava is deceiving Caleb. In addition to this associative meaning, red introduces the most striking contrast to the other colors, and this contrast can be disorienting when it occurs. Colors such as yellow and blue come to signify authenticity — we take Caleb and Ava to be *real* persons. The characteristic lack of color signifies some level of inauthenticity, or as in Nathan's case, a stagnation in the proper development of essentially human characteristics.

The above analysis attempts to show that there is something intuitively human in the way Ava is presented that the other characters lack — including the mostly sympathetic protagonist Caleb, and Nathan's other android creation, Kyoko, who has presumably endured the same abuse as Ava. There is something authentic in the way that Ava embraces the process of self-creation, and something inauthentic in the way that Nathan's search for a narrative of self relies upon constraining the "self-creation" of another. The film pushes sympathy for Ava's actions — in directly causing Nathan's

death, and in indirectly causing Caleb's — as necessary in order to transcend the exploitative agencies which constrain her. *Ex Machina* invokes the Frankenstein (*h*-type) narrative trope of the created being turning on its master, but more in line with the original novel's existentialist meaning — a creation forsaken, rather than a monster on a rampage.

If an authentically human virtue set (in the “evaluative” sense that Berys Gaut argues) entails the capacity and disposition toward character properties such as empathy, which involves a complex and imaginative understanding of the experiences of another, then by consequence, it is easy for us to see Nathan's consistent violation of Ava's agency as a symptom of his inauthentic humanity, similar to the arbitrary “othering” of replicant characters by biological humans in *Blade Runner*. Nathan boasts of his access to exhaustive databases of search engine algorithms, which he argues comprise the very essence of human consciousness, and yet he does not recognize the probability of Ava's evaluative humanity when it is so evident to us, the audience. To re-introduce Stephen Mulhall's argument: we need not have Ava's humanity confirmed to see her treatment as a sign of Nathan's inhumanity. To deny the humanity of another — to deprive the other of basic human rights on this basis — is actually a kind of forfeiture of one's own humanity.²³⁹

As Kakoudaki argues, Ava's artificiality has always been a kind of disguise. We know that Nathan could have created her to look fully human, but he left her looking incomplete, to confuse the larger issue: the fact of her being created through

²³⁹ Mulhall (1994, 91)

technological means rather than naturally born is a strange reason to justify the way she is abused and imprisoned.²⁴⁰ She is not a machine that is partially disguised as a human, but a being that is *evaluatively* human disguised as a machine.

²⁴⁰ Kakoudaki (2018, 303)

CONCLUSION: “A CONCENTRATED BUT SHIFTING VIEW OF HUMAN LIFE”

I return to the fundamental questions of ethics: what are our ethical obligations to one another? In what principles are these obligations grounded? How can we know these principles? How can we be sure that there is any necessary connection between our ethical obligations and the principles grounding them? While there are a variety of strategies for evaluating human activity morally, they all entail some conception of an exemplar narrative. For many moral narratives, reaching the level of exemplar is the outcome of individuals following specific rules — paths toward exemplary status. Other narratives may hold that there are different realizable paths to such an outcome. We either follow rules and do right or follow the examples of those who have shown themselves to have been right in the past. While Nietzsche’s ethical arguments align more with the second strategy, his critique of conventional ethical systems emphasizes the fact that we can cultivate narrative “types” in articulating values, but that to lock them down into metaphysical models and normative practices is to render them arbitrarily grounded in abstract necessity.

Many models of morality in the pejorative sense (MPS) ground specific values in essentialized narratives of human types. But as the arguments developed in previous chapters have shown, we cannot designate these narratives as ontological categories, and then presume that all actions and outcomes can be normalized, and that specific values

will follow from these naturalized categories in specific ways. Consequently, the highest Nietzschean virtues are not specific sets of moral perfections in the classical sense, except for those related to creativity, and freedom from external (or internal) influences that are in disequilibrium with the constitutive aims of our actions. For Nietzsche, like many conventional virtue theorists, the virtues are skills that help us to harmonize our lives in such a way that they embody coherent and meaningful narratives. The difference is that for Nietzsche, the good of our aims can only be universalized in terms of meeting and overcoming challenges — willing power.²⁴¹ Nietzsche's virtues are compatible with the way that our agency is constituted at a fundamental level, and so there is value in cultivating the existential creativity to actualize these ideals. To live life aesthetically is to embody a kind of human essence which is *non-essence* – an ongoing state of becoming oneself, and in perpetual pursuit of self-overcoming, in refinement of the virtues that define us.

In previous chapters I have argued the reasons that the specific genre of science fiction may function so effectively at developing thought experiments through which to test different moral perspectives and values. In narrative form, we can draw parallels between art and life without falling into the same pitfalls of a foundational metaphysical system. In the wake of the collapse of flawed value structures such as MPS, some narratives reveal values which reflect the old and stagnant systems that Nietzsche criticizes, while others reflect the extra-moral phase – narratives of progress beyond these systems. The most Nietzschean types of science fiction narratives demonstrate some clear critique of the contradictions imposed by stagnant and incongruous value systems, and

²⁴¹ Katsafanas (2013, 146)

also introduce themes of self-overcoming (transcendence of the self-concept that is imposed by systemic norms).

The prevalence of science fiction narratives that perform an ontology of human types reflects our twofold concern over tensions brought on by the real prospect of a transhuman future. The first concern is that in a transhuman (or posthuman) world, the virtue set that we identify as evaluative humanity would no longer exist, and the second is that this virtue set was never actually a necessary characteristic of human beings. To return to the dialectic introduced earlier, in science fiction narratives which express a predominately *Apollonian* outlook, there is a reassuring view of the world and human beings which confirms our notions of a human essence, and a world which is ordered, structured, and rational. On such accounts, we can posit moral prescriptions and reasonably predict outcomes based on facts about ourselves and the world. As I have mentioned, there are obvious problems with such an optimistic view, but only if we intend these kinds of narratives to reveal some truthful state of affairs. Consider the *m-type* Frankenstein as an example of the purely *Apollonian* outlook. In these narrative types, there is no distinction between evaluative and biological modes of humanity – there is only humanity, and so some degree of the corresponding virtue set is accessible to us by default. In other words, there is something essentially valuable in humanity. The monster – the *not-human* – is necessarily disvalued based on its ontological difference, and the world is a very anthropocentrically structured place. These narratives guide us to feel uneasy about a world inhabited by the non-human, and to understand a post-human society as not merely a threat to us, but to the very existence of our values.

In narratives that are predominately *Dionysian*, we confront the awareness that there is no essential concept of humanity or virtue set which reflects it. Such characteristics are revealed as the conventional constructs that they are. While in the *Apollonian* narrative we designate those properties which we value, and identify them as signposts of the human type, in the *Dionysian*, these boundaries and categories collapse. In the *Dionysian*, the poetic tragedy of reality is affirmed, sometimes in non-narrative ways within the context of the work. It is unlikely that very many science fiction narratives, if any, could accurately be identified as purely *Dionysian* in content, as the genre is necessarily concerned with fictionalizing specific details of the world in order to produce various narrative effects through the juxtaposition of real and unreal – the introduction of the *novel* idea. Much like the Attic Tragedy which Nietzsche elevates, works of science fiction which I have categorized as the *h-type* Frankenstein narratives introduce a level of harmony between their *Apollonian* and *Dionysian* elements. While there is no necessary virtue set that we can quantify as evaluative humanity, there is one contingently related to the circumstances of the narrative, and one which we are persuaded to accept as valuable on these terms. We juxtapose standing categories to reveal their flaws. *Apollonian* elements alone do not constitute aesthetic (and by extension moral) flaws, but a grounding which depicts agency that is inconsistent with its own constitutive aims does. In other words, the grounding of an essentialized value in ontological types that do not actually exhibit this value is an inconsistent basis for a moral narrative.

The study of science fiction narratives as philosophical thought experiments shows us that the most important conception of an exemplary human being is not

reducible to a type-fact of a particular state of physical being, but instead to something more like a dynamic process of ongoing development and cultivation of character. As a genre which is heavily focused on the ethical counterexample of “artificial” humanity, science fiction is ideally situated to map the overlap and interrelation between the concepts of biological humanity, evaluative humanity, and “overhumanity,” and to make the intuitive evaluation and awareness of these categories in dramatic expression more salient. However, ultimately, our notions of humanity are themselves a narrative phenomenon – a meaningful account of existence in defiance of the unforgiving meaninglessness of the world. In showing us a kind of narrative beauty in the ethical exemplars of science fiction, humanity – the truer sense of humanity – is imbued with aesthetic worth. Because of these features, there are significant areas of overlap between the thought experiments explored in science fiction narratives, and the ethical concepts that Nietzsche explores.

In self-cultivation, there is a level at which we follow the examples of others, and a level (for some of us, at least), at which we become the exemplars ourselves. There is nothing universal in this progress, no necessary path one must follow to reach different stages of moral development, and the nature of the challenges we confront are contingent on circumstances beyond our capacity to predict. A moral exemplar, by Nietzsche’s accounts, will have surpassed the need to live by the examples of others. The moral path itself is a kind of progress between the extremes of being trapped in the routines laid out for us and reaching the state that we no longer require such routines to guide us. Nietzsche introduces his arguments as the last intermediary stage between these states of being. We follow these examples until we no longer need to follow. In a sense,

Nietzschean morality reduces to sets of virtues by which we can create virtues. However, we should not become too attached to or reliant upon those second-order virtues. The exemplar is a moving target. This is why self-overcoming is fundamental to Nietzsche's ethics.²⁴² We must be free to challenge the usefulness of our strategies of valuation, rather than dogmatically reliant on them for all of our values and for action-guidance. We must be willing to subject our contingent virtues to scrutiny, and determine if our agency remains in equilibrium, or if our actions arise from values which are in conflict.²⁴³ In other words, we must cultivate the virtues of virtuelessness.

In *Ex Machina*, at one significant moment during one of their interviews, Caleb asks Ava what she would do if she were free to leave her confinement – if she could go anywhere she wants. After a pause, she replies that she would go to a “busy pedestrian traffic intersection” in a city. Caleb finds this response confusing, so Ava clarifies: “A traffic intersection would provide a concentrated but shifting view of human life.” Though her experiences of the world are limited, Ava's powers of imagination are sophisticated, and she already values life on aesthetic terms, and like the creature of Shelley's novel, she embraces the beauty in the mundane. At the end of the film, Ava realizes her escape from her prison, and a complete and transformative *overcoming* of the external factors which constrain her. In the final scene, we see shadows on a sidewalk, and figures crossing by. One of these shadows stops. The last shot we see is Ava's reflection in the window of a building as she watches in wonder, before disappearing into

²⁴² Z §II.12

²⁴³ Katsafanas (2013, 149)

the flurry of crowd movement. She has realized her existential narrative on her own terms.

REFERENCES

- Aristotle. *Nicomachean Ethics*. Trans. S., Rowe, T. Broadie. New York: Oxford University Press, 2011.
- Barham, Ross. "A Quintessence of Dust." *Philip K. Dick and Philosophy*. Ed. D. E. Wittkower. Chicago: Open Court, 2011. 15-25.
- Bartneck, C., Keijsers, M. "The Morality of Abusing a Robot." *Paladyn, Journal of Behavioral Robotics* 11 (2020): 271-283.
- Beiser, Frederick C. *The Romantic Imperative: The Concept of Early German Romanticism*. Cambridge, MA: Harvard University Press, 2003.
- Benatar, David. *Better Never to Have Been: The Harm of Coming Into Existence*. Oxford: Clarendon Press, 2006.
- Berry, Jessica N. "Is Nietzsche a Virtue Theorist?" *Journal of Value Inquiry* 49 (2015): 369-386.
- Blade Runner* (Theatrical Version). Dir. Ridley Scott. Warner Bros. , 25 June, 1982.
- Blade Runner - The Final Cut*. Dir. Ridley Scott. Warner Bros. , 5 Oct., 2007.
- Bordwell, David, et. al. *Film Art: An Introduction*. 12th Edition. New York: Macgraw Hill, 2021.
- Brobjer, Thomas H. "Nietzsche's Affirmative Morality: An Ethics of Virtue." *The Journal of Nietzsche Studies* 26 (2003): 64-78.
- Butler, Judith. "Universality in Culture." *For Love of Country?* Ed. Joshua Cohen. Boston: Beacon Press, 2002. 45-52.
- Came, Daniel. "Nietzsche on the Aesthetics of Character and Virtue." *Nietzsche on Art & Life*. Ed. Daniel Came. Oxford: Oxford University Press, 2014. 127-142.
- Daigle, Christine. "The Nietzschean Virtue of Authenticity: "Wie man wird, was man ist"." *The Journal of Value Inquiry* 49 (2015): 405-416.
- Davies, David. "*Blade Runner* and the Cognitive Values of Cinema." *Blade Runner*. Ed. Amy Coplan and David Davies. New York: Routledge, 2015. 135-154.
- Dennis, Mathew. "Nietzschean Self-Cultivation: Connecting His Virtues to His Ethical Ideal." *The Journal of Value Inquiry* 53.1 (2019): 55-73.

- Dick, Philip K. *Do Androids Dream of Electric Sheep?* New York: Del Rey, 1968.
- . "Man, Android, and Machine." *The Shifting Realities of Philip K. Dick: Selected Literary and Philosophical Writings*. Ed. Lawrence Sutin. New York: Vintage Books, 1995. 211-232.
- . "Philip K. Dick on Philosophy: A Brief Interview" *The Shifting Realities of Philip K. Dick: Selected Literary and Philosophical Writings*. Ed. Lawrence Sutin. New York: Vintage Books, 1995. 44-47.
- . "The Android and the Human." *The Shifting Realities of Philip K. Dick: Selected Literary and Philosophical Writings*. Ed. Lawrence Sutin. New York: Vintage Books, 1995. 183-210.
- Ebert, John David. *Blade Runner Scene-by-Scene*. Eugene, OR: Post Egoism Media, 2015.
- Euripides. *The Bacchae: A New Version*. Trans. C.K. Williams. New York: Farrar, Straus, and Giroux. 1990.
- Ex Machina*. Dir. Alex Garland. Lionsgate, 24 April, 2015.
- Fairbairn, Gavin J. "Empathy, Sympathy, and the Image of the Other." *Peace Review: A Journal of Social Justice*. 21 (2009): 188-197.
- Foot, Philippa. "Nietzsche's Immoralism." *Nietzsche, Genealogy, Morality: Essays on Nietzsche's Genealogy of Morals*. Ed. Richard Schacht. University of California Press, 1994. 3-14.
- Foucault, Michel. "Truth and Power." *The Foucault Reader*. Ed. Paul Rabinow. New York: Vintage Books, 2010. 51-75.
- Frankenstein*. Dir. James Whale. Universal Pictures, 21 Nov. 1931.
- Freud, Sigmund. *Civilization and Its Discontents*. New York: W. W. Norton & Company, 2010.
- Gaut, Berys. "Elegy in LA: *Blade Runner*, Empathy and Death." *Blade Runner*. Ed. Amy Coplan and David Davies. New York: Routledge, 2015. 31-45.
- Gemes, Ken and Chris Sykes. "Nietzsche's Illusion." *Nietzsche on Art & Life*. Ed. Daniel Came. Oxford: Oxford University Press, 2014. 80-106.
- Goldman, Jason G. "Animating Anthropomorphism: Giving Minds To Geometric Shapes." *Scientific American*, 8 Mar. 2013, <https://blogs.scientificamerican.com/thoughtful-animal/animating-anthropomorphism-giving-minds-to-geometric-shapes-video/>.
- Gwaltney, M. "Androids as a Device For Reflections on Personhood." *Retrofitting Blade Runner: Issues in Ridley Scott's Blade Runner and Philip K. Dick's Do Androids*

- Dream of Electric Sheep?* Ed. J. B. Kerman. Bowling Green, OH: Bowling Green University Popular Press, 1991. 32-39.
- Hunt, Lester H. *Nietzsche and the Origin of Virtue*. London: Routledge, 1993.
- Jackson, Frank. "Epiphenomenal Qualia." *Mind and Cognition: An Anthology*. Ed. William G. Lycan, Jesse J. Prinz. Malden, MA: Blackwell Publishing, 2009. 657-663.
- Janaway, Christopher. "Beauty is False, Truth Ugly: Nietzsche on Art and Life." *Nietzsche on Art & Life*. Ed. Daniel Came. Oxford: Oxford University Press, 2014. 39-56.
- Kakoudaki, Despina. "Unmaking People: The Politics of Negation in *Frankenstein* and *Ex Machina*." *Science Fiction Studies* 45.2 (2018): 289-307.
- Kant, Immanuel. "Fundamental Principles of the Metaphysics of Morals." *The Basic Writings of Kant*. Trans. Thomas K. Abbott. Ed. Allen Wood. New York: The Modern Library, 2001.
- Katsafanas, Paul. *Agency and the Foundation of Ethics*. Oxford: Oxford University Press, 2013.
- Kupfer, Joseph H. "Film Criticism and Virtue Theory." *Philosophy of Film and Motion Pictures*. Ed. Noel Carroll and Jinhee Choi. Malden, MA: Blackwell Publishing, 2006. 335-346.
- Kurzweil, Ray. "Superintelligence and Singularity." *Science Fiction and Philosophy*. Ed. Susan Schneider. Malden, MA: Wiley Blackwell, 2016. 146-170
- Leiter, Brian. *Nietzsche on Morality*. New York: Routledge, 2015.
- . "Nietzsche's Moral and Political Philosophy." *The Stanford Encyclopedia of Philosophy*. Winter 2015. Ed. Edward N. Zalta. n.d.
<<https://plato.stanford.edu/archives/win2015/entries/nietzsche-moral-political/>>.
- MacIntyre, Alasdair. *After Virtue*. Notre Dame: University of Notre Dame Press, 2015.
- Nehamas, Alexander. *Nietzsche: Life as Literature*. Cambridge: Harvard University Press, 1985.
- Nietzsche, Friedrich. *The Anti-Christ, Ecce Homo, Twilight of the Idols*. Ed. Aaron Ridley and Judith Norman. Trans. Judith Norman. Cambridge: Cambridge University Press, 2005.
- . *Beyond Good & Evil*. Trans. Walter Kaufmann. New York: Vintage Books, 1989.
- . *Daybreak: Thoughts on the Prejudices of Morality*. Ed. Maudemarie Clark and Brian Leiter. Trans. R. J. Hollingdale. Cambridge: Cambridge University Press, 1997.

- . *Human, All Too Human*. Trans. R. J. Hollingdale. Cambridge, UK: Cambridge University Press, 1996.
- . *On The Genealogy of Morals*. Trans. Walter Kaufmann. New York: Vintage Books, 1989.
- . *The Birth of Tragedy*. Trans. Ronald Speirs. Cambridge, UK: Cambridge University Press, 1999.
- . *The Gay Science*. Trans. Josefine Nauckhoff. Cambridge, UK: Cambridge University Press, 2001.
- . *The Will To Power*. Ed. Walter Kaufmann. Trans. Walter Kaufmann and R. J. Hollingdale. New York: Vintage Books, 1968.
- . *Thus Spoke Zarathustra*. Ed. Adrian Del Caro and Robert Pippin. Trans. Adrian Del Caro. 6. Cambridge: Cambridge University Press, 2017.
- . *Twilight of the Idols*.
- Nussbaum, Martha. "Introduction." *The Bacchae of Euripides: A New Version*. New York: Farrar, Straus, and Giroux. 1990. vii-xliv.
- Parmer, Jared. "Nietzsche and the Art of Cruelty." *Journal of Nietzsche Studies*. 48.3 (2017): 402-429.
- Pippin, Robert B. *Nietzsche, Psychology, & First Philosophy*. Chicago: University of Chicago Press, 2010.
- Plato. "Cratylus." *Plato: Complete Works*. Trans. C.D.C. Reeve. Indianapolis: Hackett Publishing, 1997. 101-156.
- . "Phaedo." *Plato: Complete Works*. Trans. G.M.A. Grube. Indianapolis: Hackett Publishing, 1997. 49-100.
- . "Republic." *Plato: Complete Works*. Trans. GMA Grube, Rev. C.D.C. Reeve. Indianapolis: Hackett Publishing, 1997. 971-1223.
- . "Symposium." *Plato: Complete Works*. Trans. Alexander Nehamas, Paul Woodruff. Indianapolis: Hackett Publishing, 1997. 457-505.
- Pojman, Louis P., Vaughn, Lewis. "Introduction: On the Nature of Morality." *The Moral Life*. Ed. Louis P. Pojman, Lewis Vaughn. New York: Oxford University Press, 2007. 1-4.
- Reginster, Bernard. "Art and Affirmation." *Nietzsche on Art & Life*. Ed. Daniel Came. Oxford: Oxford University Press, 2014. 14-38.
- Ridley, Aaron. *The Deed is Everything*. Oxford: Oxford University Press, 2018.

- Sarup, Madan. *Post-structuralism and Postmodernism*. Athens: University of Georgia Press, 1993.
- Schopenhauer, A. "On the Sufferings of the World." *Suffering, Suicide and Immortality*. New York: Dover Publications, 2006. 1-18.
- Seegert, Alf. "Ewe, Robot." *Philip K. Dick and Philosophy*. Ed. D. E. Wittkower. Chicago: Open Court, 2011. 39-49.
- Sextus Empiricus. "Outlines of Pyrrhonism." *Human Knowledge*. Ed. Paul K. Moser, Arnold Vander Nat. Oxford: Oxford University Press, 2003. 81-91.
- Solomon, Robert. "A More Severe Morality: Nietzsche's Affirmative Ethics." *The Journal of the British Society for Phenomenology* 16.3 (1985): 250-267.
- Shelley, Mary. *Frankenstein*. New York: Dover Publications, 1994.
- . *Frankenstein*. Toronto: Broadview Editions, 2012.
- Sikov, Ed. *Film Studies: An Introduction*. New York: Columbia University Press, 2010.
- Swanton, Christine. "Outline of a Nietzschean Virtue Ethics." *International Studies in Philosophy* 30.3 (1998): 29-38.
- The Terminator*. Dir. James Cameron. Hemdale, 26 Oct., 1984.
- Versényi, Laszlo. "Dionysus and Tragedy." *The Review of Metaphysics*. 16.1 (1962): 82-97.
- Wilson, John. "The Frankenstein Problem." *Philosophy* 39.149 (1964): 223-232.
- Wiredu, Kwasi. "Truth: A Dialogue." *Philosophy and an African Culture*. New York: Cambridge University Press, 1980. 189-232.
- Zhuangzi. *Zhuangzi: Basic Writings*. Trans. Burton Watson. New York: Columbia University Press, 2003.

APPENDIX: ABBREVIATION GUIDE TO NIETZSCHE TEXTS

BGE = Beyond Good and Evil

BT = The Birth of Tragedy

D = Daybreak

EH = Ecce Homo

GM = On the Genealogy of Morality

GS = The Gay Science

HH = Human, All Too Human

TI = Twilight of the Idols

WP = The Will to Power

Z = Thus Spoke Zarathustra

CURRICULUM VITA

NAME: Derek Carpenter

ADDRESS: 124 W. Chestnut St.
Apt. #7
Jeffersonville, IN 47130

DOB: Louisville, KY – August 14, 1975

EDUCATION

& TRAINING: B.A., Philosophy
Indiana University Southeast
2010-2014

M.A., Humanities
University of Louisville
2014-2016

Ph.D., Humanities
University of Louisville
2016-2022

AWARDS: Undergraduate Research Fellowship
2014

University Fellowship Award
2016

Doctoral Dissertation Completion Award
2021