Deriving Ethics from Action: A Nietzschean Version of Constitutivism

PAUL KATSAFANAS
Boston University

How can we justify normative claims about what there is reason to do, such as "there is reason not to lie," or "you should not murder"? Lately, a number of philosophers have argued that we can justify normative claims by deriving them from facts about the nature of action. According to *constitutivism,* action has a certain structural feature – a constitutive aim – that both constitutes events as actions and generates a standard of assessment for action. We can use this standard of assessment to derive normative claims. In short, the authority of certain normative claims arises from the bare fact that we are agents.

So, at any rate, the story goes. But should we believe the story? Although Christine Korsgaard and David Velleman have defended versions of constitutivism,[1] these theories have recently been subject to a number of powerful objections. Kieran Setiya, David Enoch, and others claim to have shown that these constitutivist theories face insurmountable difficulties; as Enoch puts it, "normativity cannot be grounded in what is constitutive of agency" (2006, 192).[2] As a result, I think it is fair to say that although constitutivism is an attractive justificatory strategy within ethics, the current versions of constitutivism are widely thought to be unsuccessful.

In this essay, I argue that the prospects for constitutivism are not so grim. I develop and defend a new version of constitutivism. I argue that by examining the structure of human motivation – by engaging in philosophical psychology – we can show that action has a constitutive aim. Moreover, I argue that this constitutive aim generates conclusions

---

[1]   See Korsgaard (1999), Korsgaard (2009), Velleman (2000), and Velleman (2009).

[2]   See, for example, Enoch (2006) and Setiya (2007).

about what there is reason to do. So we can move from philosophical psychology to ethics.

To develop the requisite account of human motivation, I turn to a largely untapped source of ideas about the relationship between agency and value: the work of Nietzsche. Nietzsche might seem to be an unpromising source for ideas conducive to the defense of constitutivism. After all, he is famously skeptical of ethical theorizing, and he flatly denies that there are any objective facts about what is valuable. However, as I will argue below, Nietzsche does offer ethical ideals of his own, and his critiques of traditional morality rely on the idea that a certain value – power, in particular – has a privileged normative status. I will argue that Nietzsche claims a privileged normative status for this value precisely because we are committed to this value merely in virtue of acting. Nietzsche's obscure claim that all actions manifest, and are to be evaluated in terms of, ''will to power'' is an attempt to move from a claim about the essential nature of action or willing to a claim about value. Thus, surprising as it may sound, I will argue that Nietzsche uses a claim about the constitutive features of action to derive a standard of success for action. His strategy therefore parallels that of the contemporary constitutivists.[3]

The Nietzschean version of constitutivism is founded upon the idea that our actions are motivated by a distinctive kind of psychological state, the *drive*. The fact that our actions are drive-motivated turns out to entail that action constitutively aims at *overcoming resistance*: whenever we will an end, we aim not merely to achieve the end, but also to encounter and overcome resistances that arise in the pursuit of the end. (This is what Nietzsche means by ''will to power.'') I will argue that we can use this feature of action to ground normative conclusions about that there is reason to do.

This essay comprises six sections. Section 1 begins with an interpretive problem: how does Nietzsche justify his own normative claims? This leads to an independent philosophical question: how can normative claims be justified? I explain constitutivism's unique and compelling answer to this question. Section 2 contends that Nietzsche's arguments concerning value are best interpreted as a version of constitutivism. Section 3 asks whether Nietzsche's view is true. I attempt to show that Nietzsche has a powerful argument for constitutivism, and that key elements of his account are supported by contemporary

---

[3]   It is worth noting that constitutivism has historical roots in the work of Plato, Aristotle, and Kant. Together with Schopenhauer, these are precisely the philosophers with whom Nietzsche was most deeply engaged. Thus, it would not be entirely surprising if Nietzsche, under the influence of these philosophers, had developed a theory of value that took the form of constitutivism.

empirical work on human motivation. Section 4 examines the normative implications of the view, explaining how the account generates claims about both what we have reason to *do* and what we have reason to *value*. Section 5 considers potential objections, and Section 6 summarizes the paper's conclusions.

1. Justifying normative claims

1.1 The interpretive problem: what justifies Nietzsche's claims about value?

When we are confronted with a normative claim, such as "you should not lie," we can ask what justifies the claim. The bulk of Nietzsche's writing is concerned with this kind of questioning. Nietzsche enjoins us to scrutinize our normative claims, asking how they arose and what compels us to obey them:

> Your judgment 'this is right' has a prehistory in your drives, inclinations, aversions, experiences, and lack of experience: you must ask '*how* did it arise?' and then '*what* is really driving me to listen to it?' (GS 335)[4]

Nietzsche's work provides a particularly vivid example of the extent to which this kind of questioning can proceed. He famously wants us to scrutinize *all* of our values, by engaging in a "revaluation of all values":

> We need a *critique* of moral values, *for once the value of these values must itself be called into question*. (GM Preface 6)

To revalue a value is to assess it, to ask whether it merits the status that we accord to it. For example, consider egalitarianism. To undertake a revaluation of egalitarianism would be to engage in a critical assessment of the value that we place on egalitarianism. We might begin by asking whether egalitarianism is *really* valuable, or whether our valuation of egalitarianism is *justified*, or whether everyone *has reason* to value egalitarianism. Once we pose these questions, we seem to need some principle, some standard, which provides an answer. For example, if we operate within a Hobbesian framework we might ask whether egalitarianism promotes rational self-interest; if we operate within a Kantian framework, we might ask whether egalitarianism

---

4   For a key to the abbreviations of Nietzsche's works, see the References.

could be autonomously willed; if we operate within a utilitarian frame-work, we might ask whether egalitarianism promotes the greatest happiness of the greatest number. In each case, the revaluation presupposes some principle, itself exempt from revaluation.

Nietzsche makes it quite clear what the terms of the revaluation are to be: he declares that the "principle" of revaluation is "will to power" (WP 391/KSA 12:2[131]). Or, as he elsewhere puts it, "what is good? Everything that heightens in human beings the feeling of power, the will to power, power itself" (A 2). So on Nietzsche's account we would reva-lue egalitarianism by inquiring into the relation between egalitarianism and power. In this sense, power has a privileged normative status: it is the one value in terms of which all other values are to be assessed.[5]

Yet the secondary literature on Nietzsche has long struggled with the question of how to justify Nietzsche's privileging of power. The problem is that Nietzsche endorses claims about value that seem to rule out the possibility of power's having a privileged status. For Nietzsche flatly denies that there are objective facts about what is valuable. He writes,

> There are altogether no moral facts. Moral judgments agree with religious ones in believing in realities where there are no realities. (TI VII.1)

> There are no moral phenomena at all, but only a moral inter-pretation of phenomena. (BGE 108)

Rather, Nietzsche tells us that all value is in some sense created by human activities. For example, he writes,

> Whatever has value in our world now does not have value in itself, according to its nature – nature is always value-less, but

---

has been *given* value at some time, as a present – and it was *we* who gave and bestowed it. (GS 301; cf. D 3)

Thus, Nietzsche accepts the following three claims about value:

1. Power has a privileged normative status.

2. There are no objective values, or there are no objective facts about what is valuable.

3. All values are created by human activities.

Unfortunately, these three claims seem to be in tension with one another. If there are no objective values, and all values are created, why should power enjoy a privileged status?

This interpretive problem reaches right to the heart of Nietzsche's work, for the bulk of Nietzsche's writing is devoted to critiquing traditional systems of values. The most common form of critique employed by Nietzsche is this: when we examine some traditional value, we see that embracing that value weakens us, or undermines our capacity for willing power. This form of critique assumes that power has a privileged status. For example, suppose that I scrutinize my valuation of egalitarianism, and discover, as Nietzsche claims, that valuing egalitarianism undermines power. Nietzsche would take this as settling the matter; egalitarianism is to be rejected. But why? Why not instead reject power? Or why not simply live with the fact that the world is inhospitable to the joint realization of two values, and strike some sort of compromise? The answer must be that power has a privileged normative status.

So we have an interpretive problem, which has bedeviled generations of commentators. The structure of the possible solutions should be clear enough. First, we could try to read Nietzsche in a way that avoids committing him to one of the claims: we could deny that he gives power a privileged status, or try to read him as claiming that power is objectively valuable, or some such. Unfortunately, this approach isn't very plausible, as each claim is amply supported by a range of textual evidence.[6] Alternatively, we could conclude that Nietzsche is simply confused, and ends up endorsing inconsistent propositions. This is not very plausible, either – these aren't arcane or deeply hidden inconsistencies, of the sort that might escape a philosopher's notice; the tensions are palpable. So what is the solution? Well, notice that the three claims would be consistent if we could somehow show that although values

---

[6] For an extended argument against these alternative interpretations, see the Appendix to Katsafanas (2008).

are created – although values spring from our own activity – there is at least one value to which we are inescapably committed.

1.2 The philosophical problem: what justifies claims about value?

Let's step back from Nietzsche's work and examine the philosophical problem that we have encountered. We want to know whether there is a way of holding that values in some sense arise from us, while at the same time maintaining that there is one value to which we are inescapably committed. To bring this problem into focus, let's examine an issue that has occupied a great deal of attention in the last few decades: the dispute between internalists and externalists about practical reason.

Take an ordinary, perfectly homely normative claim, such as "you should eat ice cream." What would make it the case that this claim is true?

Well, one familiar way of justifying this sort of claim is by linking it to the agent's desires or aims. That is, if an agent has a desire or aim that would be fulfilled or promoted by an action, then we are inclined to think that the agent has a reason to perform the action. (A defeasible reason, of course.)

Some philosophers believe that this is the *only* way to justify normative claims. This position, *internalism,* is sometimes attributed to Hume, and was popularized by Bernard Williams.[7] Internalists accept the following claim:

> Agent A has reason to ϕ iff A has a desire or aim that is suitably connected to ϕ-ing.

There are, of course, many different ways of spelling out the exact way in which the action and desire must be related, in order for the desire to count as "suitably connected" to the action. The most straightforward version claims that A has reason to ϕ iff A has a desire that is promoted or fulfilled by ϕ-ing. A more complex version, which is defended by Bernard Williams, claims that A has reason to ϕ iff A has a desire that is connected by a "sound deliberative route" to ϕ-ing.[8] However, these details won't be relevant for our purposes.

Internalism provides a straightforward and relatively uncontroversial way of justifying normative claims.[9] But it faces a potential problem,

---

[7] Williams (1981), Chapter 8.

[8] Williams (1981), Chapter 8.

[9] Internalism isn't entirely uncontroversial. See, for example, Scanlon (2000) and Dancy (2002) for arguments against the idea that desires provide agents with reasons for action. However, most philosophers believe that there are at least some cases in which desires provide reasons for action.

which can be brought out by asking what happens when an internalist attempts to explain a *moral* claim, such as "you should not murder." Moral claims have an important feature: they purport to be *non-optional* or *categorical*. That is, they purport to apply to all agents, independently of the agent's motives.

For example, consider an agent who has a strong desire to murder, and has few or no motives that would be promoted by not murdering. Despite the fact that murdering would fulfill the agent's desire, I think most of us would hold that the agent should not murder. But internalism has difficulty generating that conclusion; after all, by hypothesis the agent has no motives that would be promoted by not murdering, and has strong motives that would be promoted by murdering.

Of course, it is unlikely that very many people have motives in favor of murdering. But the example brings out a highly counterintuitive feature of internalism: if internalism is true, then it will *only be an accident* that most of us have reason not to murder. For the truth of the claim "you should not murder" will be dependent upon a contingent feature of our psychologies. If we had different motives, we would have reason to murder. And that conclusion will strike most of us as implausible.[10]

Although internalists have attempted to address this difficulty,[11] recognition of the difficulty leads some philosophers to resort to *externalism* about normative claims. Externalists hold that there can be reasons for action that do not depend on the agent's psychological makeup. In other words,

---

[10]   Strictly speaking, we should distinguish between *pro tanto* and *all things considered* reasons. Internalism can be interpreted either as a theory of all things considered reasons, or as a theory of pro tanto reasons, which would then have to be balanced in some way in order to generate an all things considered reason. Agents probably do have many pro tanto reasons to murder (for example, whenever doing so would result in the collection of an inheritance, etc.). These points won't be relevant for my purposes, though, as my point is simply that it is highly counterintuitive to suggest that merely by altering a person's motives, we could make it the case that he has an all things considered reason to murder.

[11]   Here it may be helpful to mention two quite different attempts to resolve the problem. Michael Smith endorses a version of internalism, writing "to say that we have a normative reason to ϕ in certain circumstances C is to say that we would want ourselves to ϕ in C if we were fully rational" (1994, 182). This account divorces claims about reasons from claims about the motives that the agent *actually has*, instead linking them to claims about the motives that agents *would have if they were fully rational*. Alternatively, John McDowell points out that if we accept Williams' version of internalism, which claims that an agent has a reason to ϕ iff the agent has a desire that is connected by a "sound deliberative route" to ϕ-ing, then everything will depend on what counts as a sound deliberative route. If the sound deliberative route can include the agent's appreciation of truths such as "there is reason not to murder," then the agent's reasons will not be so tightly constrained by the agent's prior psychological makeup (see "Might There be External Reasons," reprinted as Chapter Five in McDowell 2001).

> It can be true both that (i) agent A has reason to ϕ, and (ii) A has no desires or aims that are suitably connected to ϕ-ing.

Thus, a claim such as "murder is wrong" can be true independently of facts about the agent's psychology.

While externalism captures the non-optional status of moral claims, it faces several challenges. I will just mention two of them. First, there is the much-discussed problem of *practicality*. Moral claims are supposed to be capable of *moving us*. Recognizing that ϕ-ing is wrong is supposed to be capable of motivating the agent not to ϕ. But how could a claim that bears *no* relation to any of our motives possibly move us? As Williams puts it, "the whole point of external reasons statements is that they can be true independently of an agent's motivations. But nothing can explain an agent's (intentional) actions except something that motivates him so to act" (1981, 107). Williams' point is this: if the fact that *murder is wrong* is to play a role in the explanation of a person's decision not to murder, then the fact that murder is wrong must somehow figure in the etiology of the agent's action. But this suggests that, if the fact that murder is wrong is to exert a motivational influence upon the person's action, then the agent must have some motive that is suitably connected to not murdering. And this pushes us back in the direction of internalism.

Second, externalism seems susceptible to a version of Mackie's argument from queerness. Desires and aims are familiar things, so it seems easy enough to imagine that claims about reasons are claims about relations between actions and desires or aims. But what would the relata in an external reasons statement be? Are we to imagine that a claim about reasons is a claim about a relation between an action and some independently existing value? This would be odd: as Mackie puts it, "if there were objective values then they would be entities or relations of a very strange sort, utterly different than anything else in the universe" (1977, 38). For if such values existed, then it would be possible for a certain state of affairs to have "a demand for such-and-such an action somehow built into it" (1977, 40). And this, Mackie concludes, would be a decidedly odd property.

Now, I won't say much about these problems – there is a vast literature devoted to that task, and externalists have attempted to answer these challenges. The point I wish to make here is simple: both externalism and internalism have attractive features, yet incur substantial costs. Internalism grounds normative claims in familiar features of our psychologies, yet for that very reason seems incapable of generating non-optional normative claims. Externalism generates non-optional normative claims, yet encounters the problems of practicality and

queerness. It would be nice if we could preserve the attractions of these theories, while avoiding their difficulties.

Enter a third theory, which attempts to do just that: *constitutivism*. According to constitutivism, there is an element of truth in both the internalist and the externalist positions. For the constitutivist agrees with the internalist that the truth of a normative claim depends on the agent's aims, in the sense that the agent must possess a certain aim in order for the normative claim to be true. But the constitutivist traces the authority of norms to an aim that has a special status, an aim that is *constitutive* of being an agent. This constitutive aim is not optional; if you lack the aim, you are not an agent at all. So the constitutivist agrees with the internalist that practical reasons derive from the agent's aims; but the constitutivist holds that the relevant aim is one that is intrinsic to being an agent. Accordingly, the constitutivist gets the conclusion that the externalist wanted: there are non-optional reasons for acting. Put differently, there are reasons for action that arise merely from the fact that one is an agent.

So constitutivism can be viewed as an attempt to resolve the dispute between externalists and internalists about practical reason, by showing that there are reasons that arise from non-optional aims.[12]

1.3 What constitutivism is

Let's take a closer look at constitutivism. The idea behind constitutivism is straightforward: there are certain activities and event-types that *necessarily* have certain aims. There are simple examples of this phenomenon, such as the game of chess. Arguably, it is not sufficient to count as playing chess that one simply moves one's chess pieces around on the board in accordance with the rules of chess. In addition, one must aim at achieving checkmate.[13] If you do not have that aim – if you are just moving pieces, without aiming to win – then you are not really playing chess. Thus, the aim of checkmate is non-optional for chess players: if you are playing chess, then you have the aim.

Of course, the aim of checkmate can be influenced and modified by other factors. But it cannot be wholly abandoned. Consider an example. If you are playing chess with a child who is just learning the game, you may adopt the aim of letting her have a fair chance at winning.

---

[12] For this reason, constitutivism has been called both ''meta-internalism'' and ''quasi-externalism'' (Wallace 2006, 50-58; Enoch 2006, 172). Strictly speaking, constitutivism should be classified as a version of internalism. Constitutivism differs from standard internalism in that it traces norms not to a contingent desire or aim, but to an aim that is an essential feature of agency.

[13] I am simplifying a bit: one could also aim at achieving a draw.

This aim will modify the way in which you pursue the aim of achieving checkmate. For example, you may see a way to achieve checkmate, but decline to take it, in order to give the child a better chance of winning. But this kind of deviation from the activity's constitutive aim can only go so far, lest you cease to engage in the activity of playing chess. If you are not making *any* effort to achieve checkmate, then you are not *really* playing chess. Instead, you are engaged in a more complex activity, with a different constitutive aim: you are engaged in the activity of *teaching a child how to play chess*, or some such. (Notice that if you are making no effort to achieve checkmate, the child could justifiably complain that you are not really playing chess.)

Similarly, some philosophers have argued that the attitude of belief constitutively aims at truth.[14] For it seems that each instance of belief aims at truth, and aiming at truth is part of what constitutes an attitude as a case of belief. After all, if an attitude had no tendency to be responsive to indications of its truth value – if, for example, an attitude with the content *that p* persisted despite the agent's appreciation of conclusive evidence that *not p* – then the attitude would not be a belief.

Let's be more precise. We can define *constitutive aim* as follows:

(Constitutive Aim) Let A be a type of attitude or event. Let G be a goal. Then A constitutively aims at G iff

(i) each token of A aims at G, and

(ii) aiming at G is part of what constitutes an attitude or event as a token of A.

For example, suppose we let A be the attitude of *belief* and G be *truth*. Then belief has a constitutive aim of truth iff (i) each token of belief aims at truth, and (ii) aiming at truth is part of what constitutes an attitude as a belief.

We now have an account of constitutive aims. But what would follow from the fact that chess, belief, action, or some other type of attitude or event has a constitutive aim? Well, suppose we accept a relatively uncontroversial claim:

(Success) If X aims at G, then G is a standard of success for X.

For example, if chess-players aim at checkmate, then checkmate is a standard of success for chess-players. That is, we can evaluate

---

[14]   See, for example, ''On the Aim of Belief,'' reprinted in Velleman (2000).

chess-players with regard to whether they achieve checkmate. Or, if belief aims at truth, then truth is a standard of success for belief; we can evaluate beliefs with regard to whether they are true. Similarly, if action had a constitutive aim of G, then we could evaluate actions with regard to whether they achieve G.

Note that (Success) simply claims that aims generate standards of success. It applies to all aims, not just constitutive aims. Whenever you have an aim, you have a standard of success. Take our afore-mentioned chess player. Suppose she not only has the aim of check-mating her opponent, but also has the aim of enjoying her game. Then we get two standards of success: we can evaluate particular moves with regard to whether the move brings her closer to check-mate, and whether it makes the game enjoyable. These aims can interact with and modify one another: if move A would promote checkmate yet would be boring, while move B would be fascinating yet somewhat more risky, then the player may have reason to make move B. Thus, the standard induced by the constitutive aim will be one standard among many others.

So what's special about *constitutive* aims? The constitutive aim's standard of success differs from these other standards in that it is *intrinsic* to the activity in question. You can play a chess game without aiming to enjoy it, and a chess game is not necessarily defective if not enjoyed. But you can't play a chess game without aiming to achieve checkmate, so a chess game is necessarily defective if it does not achieve checkmate. Thus, the interesting feature of constitutive aims is that they generate *intrinsic* standards of success. Put differently, they generate *non-optional* standards of success.

So the important point about constitutive aims is just this: if action has a constitutive aim, then that aim will be present in every instance of action. Thus, it will give us a non-optional standard of assessment for action, a standard that applies merely in virtue of the fact that something is an action.[15]

Constitutivism therefore has several powerful advantages over other methods of justifying normative claims. Constitutivism generates non-optional normative conclusions by relying on a very spare claim about the connection between aims and standards of assessment (Success). It has the benefits of externalism, namely the capacity to generate

---

[15] Of course, the standard generated by a constitutive aim is non-optional only for those who engage in the activity that the constitutive aim governs. In the case of chess you can easily escape the standard: you can simply stop playing chess, and then the standard won't apply to you. But notice that action is crucially different: if action had a constitutive aim, you could escape the aim only if you could stop acting.

non-optional norms; but it avoids the disadvantages of externalism, namely the problems of practicality and queerness.

2. Could Nietzsche be a constitutivist?

2.1 The will to power thesis

I have just explained what constitutivism is: a method for generating evaluative standards from the claim that action has a constitutive aim. With this in mind, recall the interpretive problem with which we began. Nietzsche wants to maintain a triplet of claims about value: there are no objective values, values are created by us, and power is valuable for everyone. Notice that constitutivism has the same form: there are no objective values, rather value arises from our aims; yet there is at least one aim that is present in every agent, so there is at least one evaluative standard that holds for every agent. I submit that this structural similarity is no accident. When we examine Nietzsche's remarks on will to power, we find that he is arguing that power is the constitutive aim of action (though, of course, Nietzsche uses different terminology).

So let's look at Nietzsche's claims about will to power. What is will to power? It is important not to be misled by the surface connotations of the term "power." In ordinary discourse, the claim that people will power would suggest that people strive to dominate, tyrannize, and subjugate others. This is not what Nietzsche has in mind. Power is a term of art, for Nietzsche; he gives it a special sense.

In order to introduce this special sense, it is helpful to start with a mélange of characteristic quotations on power. Nietzsche characterizes will to power in language that seems deliberately vague; he associates power with a family of terms, such as "giving form," "expanding," "imprinting," "overcoming," "mastering," and "shaping."[16] He writes that will to power is "the will's forward thrust and again and again becoming master over that which stands in its way" (WP 696).

It is important that Nietzsche does not attribute a specific end to those who will power; the will to power can be manifest in a variety of activities that are directed at very different ends. For example, Nietzsche tells us that human beings can will power by engaging in activities as diverse as pursuing knowledge, creating art, participating in ascetic practices, and writing novels (cf. GM II.17-18 et passim).

Thus, power is not a determinate end that stands in competition with other ends that the agent might pursue. The ascetic manifests will to power by overcoming his body's resistance to suffering; the artist

---

[16]   GM II.18, GS 349, BGE 259, Z II.12, WP 696.

overcomes the difficulties inherent in turning a blank canvas into a painting; the scientist overcomes the obstacles and challenges inherent in her quest for understanding. The will to power takes manifold forms, which share only one feature: they involve a commitment to overcoming resistances or obstacles that arise in the pursuit of an end.

Thus, when Nietzsche tells us that will to power is simply "the will's forward thrust and again and again becoming master over that which stands in its way" (WP 696), when he refers to "giving form," "shaping," and so on, he means to bring out the central feature of will to power: *to will power is to aim at the activity of overcoming resistances to ends*.[17]

Now this might sound like a claim about instrumental efficacy – it might sound as if willing power is just willing to overcome whatever obstacles happen to lie between you and your end. Nietzsche means something much stronger: to will power is actively to *seek* resistances, in order to overcome them. Nietzsche makes this point in the following passages:

> The will to power can manifest itself only against resistances; therefore it seeks that which resists it... (WP 656/KSA 12.9[151])

> The will is never satisfied unless it has opponents and resistance. (WP 696)

> [Strength or power is] a desire to overcome, a desire to throw down, a desire to become master, a thirst for enemies and resistances and triumphs. (GM I.13)

In short, to will power is to aim at encountering and overcoming some obstacle or source of resistance in the pursuit of one's end.

Now we know what will to power is. But there is another important component to Nietzsche's claims about will to power: Nietzsche claims that will to power is the "essence" [*Wesen, Essenz*] of willing. There are a number of passages in the published works and unpublished notebooks that make this point. To choose just three:

> The genuinely basic drive of life [*Lebens-Grundtriebes*]... aims at *the expansion of power*... the will to power... is just the will of life [*Wille des Lebens*]. (GS 349)

---

17 My interpretation of will to power is indebted to Bernard Reginster, who argues at length for the above characterization of will to power (Reginster 2006).

> All 'purposes,' 'aims,' 'meaning' are only modes of expression and metamorphoses of one will that is inherent in all events: the will to power. To have purposes, aims, intentions, *willing* in general, is the same thing as willing to be stronger, willing to grow—and, in addition, willing the means to this. (WP 675/KSA 13.11[96])

> Everything that happens out of intentions can be reduced to the *intention of increasing power*. (KSA 12.2[88])[18]

In these quotations, Nietzsche claims that every episode of willing aims at power.

It is important to be clear on what, exactly, this means. We have just seen that willing power is aiming to encounter and overcome resistances in the pursuit of some determinate end. So power is not a first-order end; rather, willing power requires pursuing some determinate end. Thus, the will to power doctrine describes a *formal* or *structural* relation between two ends. As John Richardson puts it, the claim that we will power is not a claim about *what* we will; it is a claim about *how* we will (Richardson 1996, 21). Whenever a person wills an end, this episode of willing has a certain structure. It consists not only in the aim of achieving some end, but also in the aim of encountering and overcoming resistance in the pursuit of that end.

Given that the will to power is a formal relation, which describes the *structure* of willing, it becomes easier to see what Nietzsche means when he claims that the essence of willing is will to power. He means that power is a formal aim present in each instance of willing. Whenever an organism wills an end, the organism wills that end in a certain way: it wills to achieve that end by encountering and overcoming resistance.

2.2 Resolving the interpretive problem

Of course, the claim that we actively seek obstacles and resistances is highly counterintuitive. I will address that fact in a moment, by asking whether Nietzsche's claims about willing are defensible. First, though, I want to say a bit about what the consequences of Nietzsche's claims about power would be.

---

[18] A few more examples: "What man wants, what every smallest part of a living organism wants, is an increase in power" (WP 702/KSA 13.14[174]). "All driving force is will to power" (WP 688/KSA 13.14[121]). "Striving is nothing other than striving after power" (WP 689/KSA 13.14[81]).

What is interesting about Nietzsche's remarks on will to power is that they seem designed to show that power is the constitutive aim of willing. Suppose that one wanted to offer a constitutivist argument about willing. The first step would be to show that there is some aim that is essentially involved in each instance of willing. Notice that this is *exactly* what Nietzsche's arguments concerning the will to power are designed to establish. As we just saw, Nietzsche argues that each instance of willing aims at power. Moreover, the will to power doctrine is a claim about the essential nature of willing: it is a description of the *form* or *structure* that every episode of willing manifests. But, by the definition of Constitutive Aim, this is just to say that power is the constitutive aim of willing.

Suppose Nietzsche's arguments succeed in establishing that willing constitutively aims at power. If the constitutivist argument form were valid, then Nietzsche would be entitled to conclude that power has a privileged normative status. In drawing that conclusion, Nietzsche would *not* have to rely on the idea that power is an objective value. Rather, the argument would rely simply on the idea that insofar as an agent wills an end, the agent is committed to treating power as a standard of success for willing.

Again, this seems to be exactly what Nietzsche does conclude about power. Nietzsche denies that there are objective values, but treats power as the one standard of evaluation that readily meets challenges to its authority. And we can now see why. Nietzsche is grounding power's privileged evaluative status in an *incapacity:* it is the one value that we cannot give up, insofar as we are engaged in willing.

Surprisingly, then, Nietzsche's claims about will to power and revaluation seem to be linked by a constitutivist argument. The premises, the argument form, and the conclusion are all just what we would expect, if Nietzsche were a constitutivist.

Moreover, recall that we began examining Nietzsche's notion of will to power because Nietzsche embraces a triplet of seemingly inconsistent claims about value:

1. Power has a privileged normative status.

2. There are no objective values.

3. Values are created.

Notice that we can render these three claims consistent by interpreting power as the constitutive aim of action. If power is the constitutive aim of action, then it has a privileged normative status: it is the one

standard that is intrinsic to willing. Yet power is not an objective value; rather, it arises from the fact that our actions have a certain structure. In that sense, the value of power is created by us.

3. A Nietzschean argument for constitutivism

With that, we have solved the interpretive problem: Nietzsche's three claims about value are consistent. Values aren't out there in the world, but arise from us; yet we are inescapably committed to one value, power.

So now we know that Nietzsche has a coherent and novel view. The question is whether his view is *true*. Is it true that willing constitutively aims at power?

The hardest part of any constitutivist argument is showing that action really does have a constitutive aim. That is doubly difficult here, for Nietzsche singles out a decidedly counterintuitive aim: how could it possibly be true that in each instance of action, we aim at encountering and overcoming resistances?

As a preliminary step, it is important to note that Nietzsche is well aware that his claim will strike most readers as counterintuitive. Nietzsche is not trying to elucidate our ordinary conception of willing; rather, he is attempting to reveal the true structure of willing, which he believes has been misunderstood. This is why, in BGE 19, Nietzsche bemoans our tendency to treat willing ''as if it were the best known thing in the world.''

With that in mind, let's reconstruct Nietzsche's argument for the claim that we will power. The argument has two stages. First, Nietzsche makes a series of conceptual claims about the nature of a certain kind of motivational state, the *drive [Trieb, Instinkt]*. Given the structure of drives, it turns out that any drive-motivated action will in fact have the constitutive aim of overcoming resistance. Second, Nietzsche argues for an empirical claim, namely that *all* human actions are drive-motivated activities. If this is right, then it turns out that human action has the constitutive aim of overcoming resistance.

3.1 Outline of the will to power argument: part one

So let's start with the first stage: the conceptual claims about the nature of drives.[19] The easiest way to elucidate the nature of drives is to contrast drives with desires. Desires are typically understood as disposi-

---

[19]   Below, I summarize Nietzsche's account of drives. However, due to space constraints, I do not attempt to offer enough textual evidence to establish that my interpretation of Nietzsche's account of drives is correct. I have defended the interpretation of drives in Katsafanas (forthcoming a).

tions to bring about a goal.[20] For example, I might have a desire for food, or a desire to convince my audience of a philosophical point, or a desire to walk to my office. Drives are more complicated: drives are motivational states that have not one, but *two* quite different kinds of goals.

To clarify this point, it will be helpful to employ some terminology that Sigmund Freud introduces. Freud distinguishes between the *aim* [*Ziel*] and the *object* [*Objekt*] of the drive. The aim of the drive is its characteristic goal, in terms of which it is individuated from other drives. The aim of the sex drive is sexual activity; the aim of the aggressive drive is aggressive activity; and so on. Freud remarks, "although the ultimate aim of each drive remains unchangeable, there may yet be different paths leading to the same ultimate aim" (1957, vol. 14, 118).[21] Thus, he introduces the notion of the drive's object.

> The object of a drive is the thing in regard to which or through which the drive is able to achieve its aim. It is what is most variable about a drive and is not originally connected with it, but becomes assigned to it only in consequence of being peculiarly fitted to make satisfaction possible... It may be changed any number of times in the course of the vicissitudes which the drive undergoes during its existence... (Freud 1957, vol. 14, 118)

The *aim* of a drive is the drive's distinctive form of activity, that by means of which it is individuated. The *object* of a drive is a temporary goal upon which the drive finds expression. So, for example, while the aim of the aggressive drive is aggressive activity, the object of the aggressive drive will vary: it could be physical struggle with another individual, the participation in certain forms of sport, the infliction of physical or emotional pain on oneself (asceticism), self-repression, and so on.

---

[20] For example, Stalnaker writes, "to desire that P is to be disposed to act in ways that would tend to bring it about that P in a world in which one's beliefs, whatever they are, were true" (1984, 15). Similarly, Schroeder defines the "Standard Theory of Desire" as follows: "To desire that P is to be disposed to bring it about that P" (2004, 11). (Schroeder ultimately rejects the Standard Theory, in favor of a more complex account of desire.)

[21] In the quotations from Freud, I follow the *Standard Edition,* but make one change: whereas the *Standard Edition* mistranslates all occurrences of *Trieb* as "instinct," I translate *Trieb* as "drive." (As a host of commentators on Freud have pointed out, "instinct" has misleading connotations, and is therefore not an accurate translation of *Trieb*.)

Drives aim at their own expression, and take various objects as chance occasions for expression. The drive is not satisfied by the attainment of its objects; rather, the drive seeks these objects simply because fulfilling its aim – engaging in a form of activity – requires finding objects upon which to direct the activity. For example, to engage in aggressive activity one needs to find someone or something to be aggressive toward.

This last point will be important, so let's dwell on it for a moment. Drives don't aim at the achievement of some determinate state of affairs; rather, drives aim at the process of expression. This is a crucial difference between drives and desires, because many desires aim solely at the achievement of states of affairs: I desire that I finish my taxes, or that I take an aspirin. This is never the case with drives. Drives are constant motivational forces that incline one to engage in certain activities or processes. Drives are not satisfied by the attainment of their objects, since their objects are just chance occasions for expression. In other words, the object serves as nothing more than an opportunity for the drive to express itself. What the drive seeks is just this expression; the drive is satisfied only when being expressed, when the process that it motivates is in progress. Accordingly, an activity that is motivated by a drive does not aim at a state of affairs that puts the activity to an end; on the contrary, an activity motivated by a drive aims at the performance of the activity itself.[22]

So drives aim at expression, in the sense that they aren't satisfied by the attainment of any one determinate object; rather, they want continuous attainment of objects, continuous overcoming of resistances. We can mark this feature of drives by saying that drives are process-directed, rather than goal-directed. A goal-directed act is an act that aims at the attainment of some definite goal; thus, the act would cease, once its object is achieved. A process-directed act is an act that aims at a process; thus, there is no object the attainment of which would bring the process to an end.

With this in mind, consider again Nietzsche's claim that certain activities aim at encountering and overcoming resistance. For a goal-directed act, this claim would be absurd: if I am seeking merely to bring about some end, then it would be perverse to will resistance to that end. But for a process-directed act, things are crucially different.

---

[22] Alexander Nehamas emphasizes a related point, writing, "willing as an activity does not have an aim that is distinct from it; if it can be said to aim at anything at all, that can only be its own continuation. Willing is an activity that tends to perpetuate itself, and this tendency to the perpetuation of activity… is what Nietzsche tries to describe by the obscure and often misleading term 'will to power'" (1985, 79).

Process-directed acts do involve goals that need to be achieved, but these goals would be unimportant if divorced from the process. What matters is the process, and engaging in the process requires finding objects, resistances, challenges upon which to direct the process.

In the abstract, this may sound rather mysterious, so let's consider an example. Take the activity of running a marathon. Marathon running has two important features. First, the goal seems unimportant if divorced from the process: there is nothing particularly valuable about being in the state of having traveled twenty-six miles. But, second, the goal acquires importance when it is considered as part of the process. Running a marathon requires strenuous exertion, the overcoming of great resistance, and the experience of sometimes-intense pain. But marathon runners typically do not view these aspects of running as objectionable; on the contrary, part of the point of running a marathon is that one encounters these resistances and obstacles. In the usual case it is not that the runner values the state of having run twenty-six miles, and views the pain as a necessary, but regrettable, aspect of running. Rather, the runner values the whole activity of encountering obstacles and holding herself to a course of action despite the pain involved in doing so. This is why the runner chooses to run twenty-six miles, instead of twenty-six feet; the latter would be too easy, would not be challenging. The runner views the marathon running as valuable partly because it requires encountering and overcoming resistances and obstacles.

Thus, in the normal case, it would be distorting to view a marathon runner as aiming solely *to have run twenty-six miles*. That goal is valued only as a part of the whole process of running.

The marathon case is a good example of a process-directed act. It illustrates an important feature: continuously performing an activity entails continuously encountering and overcoming the resistances to that activity. So aiming to continuously perform an activity entails aiming to continuously encounter and overcome resistances.[23] That is, one

---

23 Nietzsche is committed to the claim that continuously performing an activity entails continuously encountering and overcoming resistance. There may appear to be counterexamples: when I loaf on the couch and watch a lowbrow sitcom on television, it may seem that I am not encountering and overcoming any resistance. However, there are resistances here, albeit of the most minimal sort: one must attend to the program, one must support oneself on the couch, one must resist competing desires that incline one to perform other actions, and so on (after all, loafing is marginally more demanding than non-action events such as sleeping). While different types of activities generate different degrees of resistance (marathoning is far more difficult than watching television), every activity generates at least *some modicum* of resistance. After all, acting is shaping a recalcitrant world: part of what it is to act is to effect a change in the world, and effecting a change in the world requires overcoming resistance. For remarks to this effect, see GM II.12, GM II.18, and WP 704/KSA 13:11[111].

does not seek merely to achieve an end state, which brings willing to an end. Rather, one seeks to continue acting, and thus seeks the conditions of continuous action.

So, while a purely goal-directed act would seek to minimize resistances, a process-directed activity involves an active desire to encounter and overcome resistances.[24] For if you aim to engage in some process, you must actually seek the objects upon which the process can be directed. This is why Nietzsche sometimes describes the will to power as "a will to overcome, a will that has in itself no end… a *processus in infinitum*, an active determining" (WP 552). Or, as he elsewhere puts it, "the will to power can manifest itself only against resistances; therefore it seeks that which resists it" (WP 656∕KSA 12.9[151]).

Let me summarize the results of this argument. First, drives are defined as motivational states that aim at their own continuous expression. Second, aiming at continuous expression entails aiming to encounter resistances to overcome. The conclusion is simple: drive-motivated activities aim at encountering and overcoming resistance. This is part of what makes an activity qualify as drive-motivated. By the definition of (Constitutive Aim), this is just to say that drive-motivated activities have the constitutive aim of overcoming resistance.

3.2 Outline of the will to power argument: part two

We have just investigated the structure of a certain kind of motivational state, the drive. We have seen that any drive-motivated activity will have the constitutive aim of overcoming resistance. Now we are faced with the obvious question: are human actions drive-motivated?

In fact, Nietzsche argues that *all* human actions are drive-motivated. If this claim is true, then it follows that human action has the constitutive aim of overcoming resistance.

Nietzsche has several different kinds of argument for the claim that human actions are drive-motivated. His most familiar form of argument is this: he will take a central type of human action which appears to have nothing to do with drives, which appears to be conditioned merely by the agent's self-conscious thought, and show that, upon closer scrutiny, the form of activity is motivated by some drive. For example, he argues that the pursuit of knowledge is motivated by what he calls the ascetic drive (cf. GM III). Taken together, these various

---

24   Resistances arise from two directions, so to speak. There are external resistances, arising from the fact that the world is recalcitrant. But there are also internal resistances, arising from competition amongst our motivational states, our drives. If each drive seeks expression, and if there are limits to the number of drives that can be expressed at any one moment, then in order to express itself a drive will have to "overpower" competing drives.

exercises might amount to an inductive proof that human actions are drive-motivated. But I won't be examining these arguments here. For I think Nietzsche has a deeper argument, which focuses on *the nature of satisfaction* (or happiness).[25] The argument claims, in essence, that the nature of human satisfaction indicates that human actions are drive-motivated.

The first claim in Nietzsche's argument is:

   (A) There is no state such that being in that state provides lasting satisfaction.

This claim is in direct opposition to a philosophical thesis that came to prominence in the work of the eighteenth- and nineteenth-century utilitarians, and that continues to exert influence in contemporary philosophical work. I will put the thesis in an unfamiliar way, and then translate it into more familiar language. Here is the unfamiliar way: Bentham and his followers thought that there were states such that human beings seek to enter into and abide in those states.

Here is a more familiar way of putting the point. Consider psychological hedonism. Psychological hedonists assume that human beings seek pleasure, where pleasure is typically conceived as an experiential *state*. Bentham wrote, "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to determine… what we shall do" (2007, 1).

Many philosophers reject psychological hedonism. However, psychological hedonism contains a thesis that is accepted even by those who allow things other than pleasure to serve as goals. This is the assumption that human beings seek *solely* to achieve goals, or to attain certain states. This assumption was explicit in eighteenth- and nineteenth-century philosophical works: we find it in Bentham, in Mill, and in Schopenhauer. For example, Schopenhauer writes, "we call [the will's] hindrance through an obstacle placed between it and its temporary goal, suffering [*Leiden*]; its attainment of the goal, on the other hand, we call satisfaction, well-being, happiness" (1969, vol. I, 309). Thus, Schopenhauer concludes that complete happiness would require "a final satisfaction of the will, after which no fresh willing would occur" (1969, vol. I, 360).

The claim, here, is that human beings seek to achieve and abide in certain states, such as the state of experiencing pleasure or the state of having one's desires fulfilled. But this very assumption – that there are states such that being in those states provides satisfaction – came under attack

---

[25]   For an interesting discussion of Nietzsche and Schopenhauer on happiness, see Reginster (2006).

by Schopenhauer. Schopenhauer argued that the attainment of a goal provides, at most, a temporary satisfaction: "with the satisfaction [of a desire], the desire and therefore the pleasure cease" (1969, vol. I, 319). For if pleasure is the state of desire-fulfillment, and desires dissipate once they are fulfilled, then pleasure is inescapably fleeting. Thus, there can be no state the occupation of which provides lasting satisfaction.

In other words: Schopenhauer argued for claim (A). And Schopenhauer famously concluded, from (A), that happiness is impossible. As he put it, "life swings to and fro like a pendulum between pain and boredom, and these two are in fact its ultimate constituents" (1969, vol. I, 312).[26]

Like Schopenhauer, Nietzsche argues for (A). But he draws a very different conclusion from the truth of (A). If we follow Schopenhauer and the early utilitarians in conceiving of happiness as an experiential *state*, then we may be led to Schopenhauer's conclusion. But, as Nietzsche puts it in one of his characteristically pithy remarks, "Man does not pursue happiness [*Glück*] – only the Englishman does that" (TI II.12). The English utilitarians' conception of happiness is defective, Nietzsche thinks. The defect is not the familiar one, not the fact that they conceive of happiness as an *experiential* state. Rather, the defect is that they conceive of happiness as a *state*.

Nietzsche rejects the state-based conception of happiness. He argues that happiness is not a *state* at all; happiness obtains when we are engaged in efficacious pursuit of a goal, not when we attain the goal. As he puts it,

> It is *not* the fulfillment [*Befriedigung*] of the will that causes pleasure [*Lust*] (I want to fight this superficial theory—the absurd psychological counterfeiting of the nearest things–), but rather the will's forward thrust and again and again becoming master over that which stands in its way. The feeling of pleasure lies precisely in the non-fulfillment [*Unbefriedigung*] of the will, in the fact that the will is never satisfied unless it has opponents and resistance. (WP 696/KSA 13.11[75])[27]

---

[26]   Here I am passing over one important component of Schopenhauer's argument, namely the claim that when we have no desires to pursue, we experience a distinctive type of pain: boredom. It is the most astonishing fact about intelligent animals that we experience boredom, and Schopenhauer is one of the few philosophers to have investigated the consequences of this fact.

[27]   Compare the following passage: "Human beings do not seek pleasure [*Lust*] and avoid displeasure [*Unlust*] … What human beings want… is an increase of power… driven by that will they seek resistance, they need something that opposes…" (WP 702/KSA 13.14[174].

Or, more succinctly:

> What is happiness?—The feeling that power *increases*—that a resistance is overcome. *Not* contentedness, but more power; *not* peace at all, but war... (A 2)

We think we are satisfied by stasis, by being in the state of having attained some goal; but in fact we are satisfied by successful activity, by actively seeking and attaining goals.[28]

So whereas Schopenhauer concludes, from (A), that satisfaction is impossible, Nietzsche concludes, from (A), that satisfaction is processual. That is,

 (B) There are processes such that engaging in those processes provides lasting satisfaction.

In other words, there are processes such that engaging in these processes provides satisfaction that lasts as long as the process lasts.

Nietzsche argues for (B) by engaging in psychological diagnoses. He develops detailed and highly insightful investigations of the vicissitudes of our actions. Some of these investigations show Nietzsche at his best and most psychologically acute, but I lack the space to detail them here.[29]

Instead, I want to take a shorter route to the conclusion that (B) is true: I will simply appeal to contemporary empirical work on human psychology.[30] For recent psychological research suggests that (B) is true. There is strong evidence suggesting that human beings are most satisfied when engaged in activities that provide them with challenges that are neither too easy nor beyond their capacities:

---

[28] Perhaps it is relevant that Dostoyevsky, whom Nietzsche referred to as "the only psychologist from whom I had something to learn" (TI IX.45), offered the same view: "Man is a frivolous and incongruous creature, and perhaps, like a chess player, loves the process of the game, not the end of it. And who knows (there is no saying with certainty), perhaps the only goal on earth to which mankind is striving lies in this incessant process of attaining, in other words, in life itself, and not in the thing to be attained" (Dostoyevsky 2006, 28).

[29] For Nietzsche's arguments in favor of (B), the following passages are especially relevant: D 60, Z Preface 3, Z II.2, Z III.5, Z IV.13, BGE 200, BGE 228, GM I.10, GM III.17, A 1-2, WP 112, WP 464, WP 1023.

[30] A version of (B) is defended by Aristotle in Book X of the *Nicomachean Ethics*. There, Aristotle argues that pleasure is something that "completes" an activity, something that occurs together with the performing of activities. Thus, rather than being something that is intelligible as a separable sensation, something that could occur apart from the performance of activities, pleasure is said to be inextricably bound to the performance of activities.

> In all the activities people in our study reported engaging in, enjoyment comes at a very specific point: whenever the opportunities for action perceived by the individual are equal to his or her capabilities... [For example,] a piece of music that is too simple relative to one's listening skills will be boring, while music that is too complex will be frustrating. Enjoyment appears at the boundary between boredom and anxiety, when the challenges are just balanced with the person's capacity to act. (Csikszentmihalyi 1990, 52)

Indeed, the most satisfying situations were found to be those in which "all of a person's relevant skills are need to cope with the challenges of a situation" (Csikszentmihalyi 1990, 53).[31]

This point can be illustrated with a mundane example from everyday life: consider the ubiquity of *games*. Take crossword puzzles. These puzzles present players with challenges that, though surmountable, require a good deal of effort. An agent who adopts the end of solving a crossword puzzle also seems to aim at encountering and overcoming resistance in the pursuit of a solution. For example, a skilled crossword puzzle player won't derive much enjoyment from solving a puzzle aimed at beginners. Part of the point of playing a game is that it is challenging, that it requires ingenuity and skill, that it presents obstacles to be overcome. One could draw similar lessons from examinations of other activities: athletes want to play teams that present challenges, not to play groups of amateurs; scientists want to discover new truths, not to go over truths already discovered.

Suppose we accept the arguments in favor of claims (A) and (B). These claims indicate that the conditions of human satisfaction have a decidedly odd structure: we desire to achieve certain ends, and are satisfied when pursuing these ends, but not when the ends have been attained. Thus, as Oscar Wilde put it, "in this world there are only two tragedies. One is not getting what one wants, and the other is getting it. The last is much the worst; the last is a real tragedy!" (1998, 37). This is paradoxical; why should it be so? What is the point of pursuing an end, if achieving the end produces at most a fleeting satisfaction? Why would we take satisfaction in this process?

Well, Nietzsche's work provides an explanation: claim (B) indicates that we aim not only to achieve ends, but also simply to be active. We aim to engage in the process of encountering and overcoming resistances in the pursuit of ends. Activities such as marathon running and

---

[31] For introductions to the psychological work on satisfaction, see for example Bradburn (1969), Csikszentmihalyi (1988) and (1990), and Deci and Ryan (1985).

puzzle solving are exemplary: they involve aiming to attain a state that will provide at most a temporary satisfaction, a state that can seem valueless. Who cares whether I have run twenty-six miles, or whether I have filled in the blanks in a crossword puzzle? But the point of these activities isn't the end – it is the process, the activity itself. And Nietzsche has a straightforward explanation for this otherwise puzzling feature of human agency: the explanation is simply that we are motivated by drives. For recall that drives aim at their own expression, at manifesting their characteristic form of activity, and take objects only as chance occasions for this expression. The attainment of objects doesn't satisfy the drive; only the performance of the drive's characteristic form of activity satisfies the drive. If our actions were drive-motivated, (A) and (B) would follow. If our actions were not drive-motivated, (A) and (B) would be puzzling. Therefore, Nietzsche argues, the best explanation for (A) and (B) is that our actions are drive-motivated.

Of course, this falls short of a deductive proof that our actions are drive-motivated. It just gives us some reason for believing that they are. As I mentioned above, Nietzsche elsewhere offers other arguments, which bolster his case for the claim that our actions are drive-motivated. But for now, I want to rest content with a modest claim: we have some reason for believing that human actions are drive-motivated.

4. Normative implications of the theory

We now have the outline of Nietzsche's argument before us. In essence, it can be summarized in just two claims. First, drive-motivated actions constitutively aim at overcoming resistance. Second, all human actions are drive-motivated. It follows that human actions constitutively aim at overcoming resistance. Or, to put this back into Nietzschean terminology: the constitutive aim of willing is power. In the concluding sections of this essay, I will briefly examine the normative implications of this view.

4.1 How the account generates claims about what we have reason to do

If the constitutive aim of action is encountering and overcoming resistance, then actions are successful to the extent that they encounter and overcome resistance. We can use this standard of success to generate ethical claims in two distinct ways.

First, the theory entails that we have reason to seek those actions that afford resistance. So, for example, if I aim to write a philosophical essay, I have reason to engage in a challenging version of that task, by choosing a topic that presents me with difficulties. (In a moment I will

introduce a complication: the degree of resistance encountered and overcome is not the *only* standard of success.)

In order to apply this prescription to a determinate case, we must take account of two facts. First, which actions an agent has reason to perform depends on facts about her capacities and psychological makeup. Most straightforwardly, an individual has reason to pursue those activities that afford resistances that are difficult to overcome, but not so difficult as to be impossible for the agent to overcome. If Sally is a brilliant novelist, and Bill can barely string together a coherent sentence, then Sally has reason to pursue the task of writing a great novel, whereas Bill does not.

Second, the results depend on facts about one's environment. The particular courses of action that constitute the highest realizations of will to power differ according to the circumstances in which an agent finds himself. During some historical epochs, the opportunities for the highest expressions of will to power involve distasteful events. In ancient societies, Nietzsche likes to remind us, the greatest expressions of will to power occurred in acts of physical violence, subjugations of one's enemies, and the like. It is important not to sanitize Nietzsche: as he was well aware, it *is* an implication of his view that, given certain pre-societal conditions, the flourishing human being will be the one who engages in the most horrific acts of oppression and subjugation.

However, Nietzsche's writings do not urge contemporary individuals to engage in these kinds of activities. Rather, Nietzsche enjoins us to undertake activities such as artistic creation, the pursuit of self-understanding, and the willingness to subject oneself to suffering in order to achieve goals. For this reason, the figures whom Nietzsche most often presents as ethical exemplars are Goethe, Beethoven, and – rather immodestly – Nietzsche himself. So Nietzsche seems to believe that in our historical circumstances, his theory recommends intellectual, artistic, and creative endeavors, rather than brute assertion of physical might or dominance.

To see why Nietzsche presents his theory as entailing these results, we must turn to a more general point. Consider the claim that action's standard of success is the amount of resistance that an agent encounters and overcomes. This doesn't seem to be a good candidate for the *only* standard that informs our deliberations. If it were, we would all have reason to perform seemingly pointless, yet immensely difficult actions, such as sticking our arms in fires or chopping off limbs. Nietzsche certainly doesn't recommend those actions; he doesn't think we select actions *merely* in terms of how much resistance they provide. Rather, the degree of resistance afforded by an action seems to be *one*

important factor among many others. In order to understand this point, we will need to examine a second aspect of Nietzsche's view.

4.2 How the account generates claims about what we have reason to value

Values pervade our world. Nietzsche claims that our experience of the world is value-laden: "There is no doubt that all sense perceptions are wholly permeated with value-judgments" (WP 505). So, too, deliberation involves the deployment of values; an agent who is deliberating about what to do will make her decision, in part, by considering her own values. Nietzsche clearly does not think that power is the *only* value influencing perception and deliberation: "the extent of moral evaluations: they play a part in almost every sense impression. Our world is *colored* by them" (WP 260). Nor does Nietzsche think it *should* be: on the contrary, he enjoins us to create new values.

Power is one value, yet is not the only value. This raises the question of what relationship these values have with one another. Interestingly, Nietzsche frequently presents will to power as a "principle of revaluation." He argues that will to power generates a standard according to which we can assess not only actions, but also *values* (cf. A 2, A6, WP 391, WP 674). Thus, I take it that Nietzsche operates with the following picture: facts about which actions we have reason to perform are determined not only by the degree of resistance that the potential actions afford, but also by facts about the relationship between potential actions and our values; however, these values must in some way be vetted by the standard of will to power. In the following sections, I will elaborate upon and defend this reading.[32]

Let's begin by asking what a value is. For Nietzsche, values are simply the objects of *valuings*: "only through valuing is there value" (Z I.15; cf. GS 301).[33] In other words, to say that X is a value is just to say that we value X. Valuing X involves having a certain kind of affective response to X, and making certain kinds of judgments about X. The particular kinds of affective responses and judgments will depend on the type of object or state of affairs that is being valued. Take compassion, for example. Valuing compassion involves judging that one has reason to act compassionately, praising compassionate behavior, striving to be compassionate, having feelings of guilt when one fails to be compassionate, and so forth. Thus, when Nietzsche speaks of

---

[32]   It is worth noting that Nietzsche's theory departs from other constitutivist theories on this point. Velleman, for example, does not present his theory as having direct implications about the values that we should embrace.

[33]   Richardson emphasizes this point (2004, Chapter 2).

compassion as one of our values, he simply means that we make certain kinds of judgments about compassion (e.g., that there is reason to be compassionate), that we have certain kinds of affects directed at compassionate behavior, and so on.[34]

These spare remarks about value are sufficient for our purposes. The only point that will matter is this: an agent's values constrain and influence his behavior, in part by influencing his judgments about reasons.[35] Accordingly, the agent's values will interact with action's constitutive standard. The constitutive standard generates claims about which acts there is reason to perform; so, too, do the agent's values. Given that claims about reasons arise from these disparate sources, there is a potential for conflict.

For an example, consider someone who values a form of complacency. This individual believes that it is valuable to be content with what one already has: one should not seek further accomplishments. This value clearly conflicts with will to power. As the prior sections argued, will to power commits us to aiming at resistances and challenges. So we have a straightforward conflict: valuing complacency involves judging that there is reason not to confront challenges; valuing power involves judging that there is reason to confront challenges. If an agent harbors both of these values, he will be in a state of conflict, endorsing contradictory propositions about how to act.

What does this tell us about the value of complacency? It is clear enough that, in presenting will to power as the standard of revaluation, Nietzsche wants us to reject any value that generates conflicts of this form. I take it that this is what Nietzsche has in mind when he writes that the "standard by which the value of moral evaluation is to be determined" is "will to power" (WP 391). In this passage, and a number of other sections, Nietzsche suggests that any value that conflicts with will to power should be rejected.[36]

However, this does raise a question. Suppose it is true that the reasons generated by valuing complacency conflict with the reasons arising from the constitutive standard. We could eliminate this conflict in two

---

[34]   Clark and Dudrick (2007) and Poellner (2007) provide nuanced analyses of the particular forms of affects and types of judgments that constitute valuings. For our purposes, the details will not be relevant.

[35]   This point is brought out in a recent discussion from T.M. Scanlon: "To value something is to take oneself to have reasons for holding certain positive attitudes toward it and for acting in certain ways with regard to it. Exactly what these reasons are, and what actions and attitudes they support, will be different in different cases" (2000, 95). As Scanlon points out, an agent's values influence the reasons that the agent takes herself to have.

[36]   For additional examples, see A 2 and GM Preface.

ways: we could cease to value either complacency or power. So why does Nietzsche enjoin us to reject complacency rather than power?

Here, Nietzsche relies on the inescapability of the constitutive standard. The values arising from our agential nature cannot be reassessed or altered; as the prior sections have argued, we are inescapably committed to valuing power. However, Nietzsche maintains that other values *can* be reassessed and altered: there are many different, mutually incompatible sets of values that we can embrace. For example, we could – and many do – regard complacency as not valuable, or even as disvaluable. So Nietzsche's point is simple: when there is a conflict between the constitutive standard and some other value, the only way in which we can alleviate the conflict is by modifying the other value.[37]

We are now in a position to see an important difference between Nietzsche's theory and more familiar ethical views. Nietzsche does not believe that in order for a value to be justified or legitimate, it must be *derived* from the will to power. To elucidate this point, it will be helpful to contrast Nietzsche's view with more familiar ethical views. Two of the most prominent ethical views, Kantianism and Utilitarianism, have a foundationalist structure: they are committed to the idea that there is *one* source of ethical norms. Kantians claim that we can derive all ethical content from one principle, the Categorical Imperative. Likewise, Utilitarians believe that we can derive all ethical content from a principle enjoining us to maximize aggregate utility. On these views, normative claims are legitimate only insofar as they follow from the respective foundational principle. For example, consider a normative claim such as "lying is wrong." On a Kantian view, this claim will be true if and only if it follows from the Categorical Imperative; on a Utilitarian view, it will be true only if acceptance of it maximizes aggregate utility.[38]

If Nietzsche's view were intended to function in an analogous fashion, then all legitimate normative claims would be derived from facts about the will to power. However, this approach would give us a set of norms so severely attenuated as to defy description as an ethic. For it should be clear enough that we are *not* going to be able to derive

---

[37] I do not mean to suggest that changing one's values is always easy. Changing one's values is typically a gradual, aggregative process, and can involve great struggles. Consider the hackneyed example of a formerly religious individual, who used to believe that sex was disvaluable, but now regards it as valuable. We can easily imagine that in order genuinely to rid himself of the old evaluation, he will have to struggle with residual feelings of guilt, shame, and so forth.

[38] I am being deliberately vague here, as the details will vary depending upon the particular version of utilitarianism that we embrace.

claims such as "lying is wrong" and "murder is wrong" from facts about will to power; there is no way of moving from the idea that we aim to encounter and overcome resistance to the idea that we should not lie, or that we should not murder. On the contrary, lying and murdering are ways – possibly quite good ways – of willing power.

Fortunately, Nietzsche's will to power doctrine is *not* intended to function as a foundational principle from which we derive all other normative claims. Rather, as this section has explained, will to power is intended to serve as a "principle of revaluation." That is, the will to power generates a standard in terms of which we are to assess all other values. So Nietzsche grounds one normative principle in facts about our agential nature, and uses this principle *not to derive, but to assess,* the other values that we embrace.

In sum, Nietzsche holds that values are legitimate insofar as they do *not* generate conflicts with will to power.[39] This allows us to see why Nietzsche's view does not simply generate a crude injunction to maximize resistance. The degree of resistance that an action affords is *one* standard of assessment, but not the *only* standard. In addition to evaluating potential actions with regard to the degree of resistance they afford, we should evaluate them with other values that we embrace.

This explains why Nietzsche's theory does not entail that we all have reason to perform every activity that generates great resistance, such as murdering, sticking our hands in fires, and so forth. These pursuits would conflict with the other values that we embrace. For example, while murdering would engender resistances to overcome, it would conflict with the value that we place on human life. While sticking one's hand in the fire would generate great resistance, it would conflict with the value that we tend to place on self-preservation. On the other hand, activities such as the pursuit of intellectual endeavors, would not conflict with, and in some cases will be supported by, our other values. Thus, there is more reason to perform these activities.

---

[39] If we are committed to willing power, shouldn't we *welcome* the resistance that obtains when we embrace values that conflict with will to power? That is, wouldn't embracing conflicting values be one way of willing power? If so, this would undermine Nietzsche's view. However, I think we can see that the answer is no. First, Nietzsche does not claim that we have reason to seek *all* forms of resistance; rather, we have reason to seek the forms of resistance that are related to the activities we are pursuing. Second, when I say that values *conflict* with will to power, I do not mean that they engender resistances. Rather, I mean that they generate contradictory claims about what there is reason to do. This isn't a conflict that one can overcome or eliminate, anymore than one could overcome or eliminate the contradictory claims generated by believing both that $1 + 1 = 2$ and that $1 + 1 = 3$.

4.3 Summary of the ethical implications of Nietzsche's view

The prior sections have shown that Nietzsche does not intend will to power to serve as the *only* legitimate value. Although will to power has a privileged normative status, it is one value among many others. Thus, the ethical implications of Nietzsche's view depend upon what *other* values one embraces. An act such as being a murderer is quite challenging, and in that respect performing the act would fulfill the agent's aim of encountering and overcoming resistance. But, of course, the fact that murder is one way of fulfilling the will to power does not imply that we should murder. After all, we accept the following normative claims: murder is wrong, human life is valuable, gratuitous infliction of suffering upon others is wrong, etc. These normative claims entail that murder is disvaluable. So, on Nietzsche's view, we would have to ask whether there are other acts that produce just as much resistance as murder, but do not violate our evaluative commitments. And, as Nietzsche makes clear, there *are*. Nietzsche's favorite examples of such acts are, as we have seen, artistic and intellectual endeavors.[40]

In sum, then, Nietzsche's ethical theory requires that we assess potential acts along two dimensions:

(1) The extent to which the act would present opportunities for encountering and overcoming resistance.

(2) Whether the act is permitted, recommended, or forbidden by the other values that we embrace.

Of course, as Section 4.2 pointed out, we must also assess our other values in terms of power. So we need a third dimension:

(3) The extent to which the values cited in (2) are compatible with the will to power.

Only by answering these three questions can we generate normative results.

---

[40] It is worth noting that Nietzsche sometimes makes this point completely explicit. For example, he writes, "It goes without saying that I do not deny – unless I am a fool – that many actions called immoral ought to be avoided and resisted, or that many actions called moral ought to be done and encouraged – but I think the one should be encouraged and the other avoided *for other reasons than hitherto*" (D 103). On Nietzsche's view, we have misunderstood the basis of ethical claims; we thought they had deeper foundations than they actually do. But this does not entail that, in every case in which traditional morality tells us *not to A*, Nietzsche thinks that *A-ing* is permissible.

So Nietzsche's evaluative framework functions in the following way. First, an agent undertakes a revaluation of values, assessing her values in light of power. She sheds some values and embraces others. She then uses this new set of values, together with the valuation of power, to determine what she should do.

As a result, Nietzsche's ethic is always historically situated: it tells us how to *go on*, not how to *begin*. In other words, Nietzschean revaluation always begins in some determinate cultural setting. An agent starts with an array of accepted values. The agent diagnoses these values in terms of will to power, abandoning some, preserving others. Where the agent ends up will depend, in part, on where she began.

5. Objections and replies

5.1 Are inescapable aims normative?

Nietzsche's theory relies upon the idea that inescapable aims generate reasons. But one might object to this claim. Consider a potential counterexample: historically, a number of ethical views have held both that (a) the sex drive is ineradicable, and (b) we have reason *not* to act on the sex drive. Here, an inescapable aim is taken not to provide reasons. Analogously, even if it were true that we inescapably aim at power, couldn't we reject the claim that we have reason to seek power?[41]

To answer this objection, let's distinguish two senses in which an aim can be inescapable:

(1) Human beings have an ineradicable desire for G.

(2) Human beings have an aim of G that is ineradicably present in every action.

Desires for food and (arguably) sex are good examples of (1): it is unlikely that an agent could eliminate these desires. However, notice that these desires are not examples of (2): it is not the case that *every* action we perform aims at food or sex. Constitutive aims, by contrast, would be examples of (2).

With this distinction at hand, we can reexamine the objection. Desires of type (1), although inescapable in one sense, are escapable in another. Although we cannot eliminate the desires for food or sex,

---

41 Thanks to an anonymous reviewer for raising this point and suggesting the potential counterexample.

we can strive *not to act on them*. When sexual urges strike, we can strive to resist them. Thus, it is at least conceivable—though no doubt unlikely in practice—that an agent could have an ineradicable desire for G without ever performing an action that aims at G. An ethical theory would not be making any *logical* error in promoting this.

However, the constitutive aim is crucially different: enjoining an individual not to act on it would be confused. If action constitutively aims at G, then *every* action that an agent performs will aim at G. So an injunction not to act on the aim would make no sense.

With this in mind, consider the claim that I labeled Success (see Section 1.3): if X aims at G, then G is a standard of success for X. Even the most minimal accounts of practical reason, including most variants of the Humean account, accept this claim. If we grant Nietzsche this spare claim about reasons, it follows that every aim is reason-providing. If an aim is ineradicably present in action, then every agent will have a reason to fulfill it. Now, in the case of sex and food, one would cease to have any reason whatsoever to engage in these activities, if one could eliminate the aim's presence in one's actions. But with the constitutive aim, this is not possible, so every agent has reason to fulfill it.[42]

5.2 Does the constitutive aim generate pro tanto reasons or all-things-considered reasons?

Suppose the arguments in the prior section establish that the constitutive aim generates reasons. Still, an objection might arise. The reasons generated by the constitutive aim are merely *pro tanto* reasons; thus, they could be overridden by the reasons generated by other aims. But

---

[42]   Enoch (2006) has a more general objection to constitutivism: even if aiming at G is constitutive of action, we could still ask why we should perform actions—why not, instead, perform ''schmactions,'' which are defined as ''nonaction events that are very similar to actions but that lack the aim (constitutive of actions but not of schmactions) of [G]'' (2006, 179)? This objection might seem to pose problems for Korsgaard's and Velleman's constitutivist theories, because those theories begin with claims about the *concept* of action, and thus might seem to leave open the possibility of an agent's simply declining to perform behaviors that meet the conditions for being labeled ''actions.'' (Thus, Enoch claims that constitutivists seem to believe ''that the threat that your inner (and outer) states will fail to deserve folktheoretical terms (such as ''action'') is indeed a threat that will strike terror in the hearts of the wicked'' [2006, 179-80].) However, Enoch's objection gets no purchase with the Nietzschean theory. According to Nietzsche, our drives induce in all of our actions an aim of power. This aim is inescapable: if Nietzsche is correct about the philosophical psychology, then we cannot avoid acting on it. So Nietzsche would argue that performing ''schmactions''—i.e., behaviors that do not aim at power—is impossible.

doesn't Nietzsche assume that the reasons generated by will to power should be overriding in cases of conflict?[43]

In fact, Nietzsche is only committed to the claim that the constitutive aim generates pro tanto reasons. In this respect, the constitutive aim is on par with all other aims. Although this fact might seem to vitiate Nietzsche's normative conclusions, I will argue that it does not.

Let's return to the chess example from Section 1.3. There, I pointed out that the reasons generated by the constitutive aim (achieving checkmate) may conflict with the reasons generated by other aims, such as the aim of teaching a child how to play chess. Consider such a case: I see that I could achieve checkmate in one move, but I believe that doing so would not be edifying for the child. I therefore consider not checkmating him, and instead moving in a way that will prolong the game. It is natural to assume that in some such cases the balance of reasons will be in favor of *not* checkmating my opponent. So here we have a pro tanto reason to checkmate, a pro tanto reason to prolong the game, and (let's stipulate) an all-things-considered reason to prolong the game. As this example indicates, chess players do not always have an overriding reason to fulfill the constitutive aim of checkmate.

Just as the constitutivist about chess will not be bothered by the fact that chess players might have reason not to pursue checkmate in a given move, so too the constitutivist about action will not be bothered by the fact that agents sometimes have reason not to perform the action that best fulfills the constitutive aim. Consider an example: suppose I can either loaf about and watch television (call this action A), or I can continue working on a difficult philosophical problem (call this action B). Action B generates far more resistance, and thus better satisfies will to power. However, having worked on the problem for several hours, I am tired and strongly motivated to take a break. These motives incline me toward A-ing. In some cases of this form, the balance of reasons may favor A-ing.

Why aren't these results troubling for the constitutivist? Although in certain circumstances there is reason not to perform the action that best fulfills the constitutive aim, things typically change when we consider longer-term actions and projects. To see why, let's alter the above example. Rather than temporarily departing from my philosophical projects in order to relax, I consider wholly abandoning these projects and spending my life doing nothing more than loafing about and watching television. So we have a series of loafing actions, A_1 to A_n, making up a larger action, A. And, as an alternative, we have a series of working actions, B_1 to B_n, making up a larger action, B. Nietzsche's

---

[43]   Thanks to an anonymous reviewer for pressing me on this point.

claim is that for each A_i, I have a will-to-power-derived pro tanto reason *not* to engage in it, whereas for each B_i, I have a will-to-power-derived pro tanto reason to engage in it. I also have other motives, so for some—but usually not all—of the A_i's, I may have a motive and thus a reason to engage in it. However, if each reason is weighted equally, I will typically have more reason to B than to A.

As this example indicates, in the long run we typically have more reason to perform the action that better fulfills the constitutive aim. It is not the *weight* of the constitutive-aim-derived reasons that generates this normative conclusion; it is their *ubiquity* and *pervasiveness*.[44]

However, consider the following potential objection: an agent could have a motive—or a set of motives—that is directly opposed to, and as ubiquitous as, the will to power. In such a case, the balance of reasons might consistently come out *against* performing the action that best fulfills will to power. Wouldn't this undermine Nietzsche's conclusions?

Nietzsche could respond to such an objection as follows: while this is conceivable, it is highly unlikely. (Notice that Nietzsche's theory, being empirically grounded in facts about human psychology, needn't commit itself to the claim that such a scenario is impossible or incoherent.) The constitutive aim typically does not *compete with* an agent's first-order motives. Rather, it modifies the manner in which these motives are pursued. Consider two examples. The desire for food could be easily satisfied by most individuals with sufficient resources – we could simply eat the same thing every day. However, very few people do this: almost no one wants to eat just anything. So, too, the desire for sex would be fairly easy to satisfy, except that almost no one desires to have sex with just anyone. Agents acting on the desires for food or sex are pursuing these goals in a way that introduces resistances – in other words, they are willing power – but they typically don't experience a conflict between, say, the (easily fulfilled) desire to have sex with just anyone and the (more difficult to fulfill) desire to have sex with a certain individual (or with an individual who has particular characteristics). *For agents typically don't have the former desire.* The will to power shapes our motives, in such a way that our motives come to express not just a desire for their determinate goals, but a desire to achieve those goals in a way that introduces resistances.

This makes long-term conflicts between will to power and other motives unlikely: our other motives are likely to *reinforce,* rather than compete with, the will to power. Accordingly, in many cases our

---

44   Velleman endorses a similar claim. He writes that the fact that an aim is constitutive of action "does not entail that it was of greater importance or influence in [agents'] lives [than their other aims]; on the contrary, it was of minor importance and influence—minor but also persistent and pervasive" (2009, 157).

motives will generate pro tanto reasons to engage in the actions that better fulfill will to power, and it is unlikely that there would be a persistent conflict between will to power and our other motives. Thus, weighting all aim-derived reasons equally, the balance of reasons is likely to come out in favor of those actions that better fulfill will to power.[45]

In sum, the constitutive aim functions not by generating overriding reasons in particular cases of action, but rather by weighting the balance of reasons in favor of the projects and long-term actions that best fulfill will to power.[46] Accordingly, there is no guarantee that in all cases of conflict we will have more reason to perform the action that better fulfills will to power: Nietzsche writes that "everything unconditional belongs in pathology," and this would surely include an unconditional requirement to maximize the fulfillment of will to power (BGE 154). However, in the long run, the balance of reasons typically favors actions that better fulfill will to power.[47] Thus, although the constitutive aim generates only pro tanto reasons, the ubiquity of these pro tanto reasons engenders substantive normative conclusions.

5.3 Are Nietzsche's critiques of more familiar ethical theories founded on his will to power thesis?

Many of Nietzsche's writings purport to undermine familiar ethical views, including in particular Christian morality and Kantian ethics (see, for example, A and EH IV). One might wonder whether Nietzsche's arguments against these views are based upon his will to power theory. If so, a problem seems to arise: the conflict between these views and the will to power thesis isn't obvious. Consider two central examples. Certain Christian moral views treat striving and perseverance as important virtues. Moreover, Kantianism isn't clearly inconsistent with the philosophical psychology that Nietzsche embraces: it seems that Kant could admit that we will power.[48]

In response, I think it would be too quick to read *all* of Nietzsche's critiques as based directly upon the will to power thesis. To be sure,

---

45 For a fascinating reflection on the way that easily fulfilled desires tend to transform into desires that engender new resistances, see Gass (1971).

46 The same point applies to *evaluative commitments,* which structure large portions of our lives. See Section 4.2.

47 Incidentally, this explains why Nietzsche's ethical writings have a peculiar feature. Unlike many other writers interested in ethics, Nietzsche very rarely focuses on the rightness or wrongness of particular, discrete actions. Rather, his writings tend to focus on long-term actions, projects, and patterns of behavior. We can now see why: these longer-term actions are where the import of the constitutive aim becomes apparent.

48 Thanks to an anonymous reviewer for asking me to address these points.

Nietzsche sometimes does argue against an ethical theory simply by claiming that it conflicts with will to power. This is the case with Christianity, as I explain below. However, at other times he critiques ethical theories on grounds independent of the will to power. This is the case with Nietzsche's arguments against Kant.

Start with Christianity. One of Nietzsche's primary objections to Christianity is that it has a state-based conception of striving: while it valorizes striving, it views striving as justified by the goal toward which it is directed—salvation, which is to be realized in an afterlife.[49] The will to power thesis, however, entails that any state-based interpretation of striving is misguided. We do not strive primarily to reach the state of having attained our goals; rather, we strive to be engaged in the very process of attaining goals. Thus, Nietzsche claims that the Christian conception of striving mistakenly views striving as something that must be justified by a state to be attained in the future, whereas the will to power thesis denies precisely this point.[50] *This* is how the will to power thesis undermines certain elements of Christian morality.[51]

With respect to Kant, the conflict is less direct. Presumably, Kantians could accept the claim that will to power is a constitutive aim of action. This would add a new normative standard to Kant's theory: in addition to acting on the Categorical Imperative, one would have reason to engage in those actions that engender obstacles and resistances.[52] Although Kantians might accept this hybrid theory, Nietzsche would reject it. For Nietzsche repeatedly claims that Kant's argument for the Categorical Imperative is unsuccessful.[53] Accordingly, Nietzsche

---

[49]   For example, consider the following passage from a contemporary of Nietzsche's, the German theologian Bernhard Weiss: ''the proclamation of the salvation given in Christ stirs up zeal to strive after Christian virtue'' (1883, §128).

[50]   See especially GM III.11, BGE 200, TI III.1, and EH IV.8. Reginster (2006) provides an excellent discussion of this point in relation to Christianity. In addition, Nietzsche has a second and much simpler objection the Christian interpretation of striving: he denies that there is an afterlife.

[51]   This point applies to other Christian values, as well. For example, Nietzsche argues that the will to power thesis conflicts with the Christian valuation of compassion. For a helpful discussion, see Chapter Four of Reginster (2006).

[52]   Kant himself might find this view agreeable: witness Kant's claim that "virtue is the strength of a human being's maxims in fulfilling his duty. – Strength of any kind can be recognized *only by the obstacles it can overcome…*" (*Metaphysics of Morals* 6:394, emphasis added). However, Kant goes on to suggest that the relevant obstacles are one's own "natural inclinations" — a claim that Nietzsche would reject.

[53]   See especially BGE 5, BGE 186, BGE 211, GS 335, and A 11-12. Nietzsche's arguments against Kant are complex, and I lack the space to reconstruct them here. I discuss Nietzsche's critiques of Kant's attempt to establish the authority of the Categorical Imperative in Katsafanas (forthcoming b).

maintains that we have no reason to govern our actions with the Categorical Imperative. Thus, while there is no direct conflict between Kant's Categorical Imperative and Nietzsche's will to power thesis, Nietzsche has independent arguments for his claim that Kantian ethics is unsuccessful. (Of course, an adequate treatment of this point would necessitate a paper of its own.)

6. Concluding remarks

Interpreting power as the constitutive aim of action resolves a long-standing interpretive puzzle in Nietzsche scholarship, by enabling us to see how Nietzsche can consistently maintain the following three claims: (1) there are no objective values; (2) values are created; (3) power has a privileged normative status. According to the interpretation that I have advanced, power is not an objective value, in the sense that it would not have value independently of a particular feature of human activities. Rather, we are committed to valuing power merely in virtue of acting, because power is the constitutive aim of action. Thus, power has a privileged normative status. Moreover, there is a sense in which the value of power is created by human activity: the structure of our own actions commits us to valuing power. Nietzsche's three claims about power and value are therefore consistent.

I have argued that once the structure of Nietzsche's will to power theory is rendered explicit, the theory becomes compelling. The theory is grounded in facts about the nature of human motivation, facts that are not only philosophically plausible, but are also supported by empirical research on human agency. Thus, Nietzsche provides us with a successful version of constitutivism.[54]

References

Reference edition of Nietzsche's works:

Friedrich Nietzsche: Sämtliche Werke, Kritische Studienausgabe in 15 Bänden, herausgegeben von G. Colli und M. Montinari (Berlin: Walter de Gruyter, 1967–1977).

List of Abbreviations of Nietzsche's Works:

A   *The Antichrist*, trans. W. Kaufmann (Viking, 1954)

BGE *Beyond Good and Evil*, trans. Kaufmann (Modern Library, 1968)

D   *Daybreak*, trans. R.J. Hollingdale (Cambridge University Press, 1982)

EH   *Ecce Homo*, trans. Kaufmann (Modern Library, 1968)

GM *On the Genealogy of Morality*, trans. Kaufmann (Modern Library, 1968)

GS   *The Gay Science*, trans. Kaufmann (Vintage, 1974)

KSA *Kritische Studienausgabe*

TI   *Twilight of the Idols*, trans. Kaufmann (Viking, 1954)

WP *The Will to Power,* trans. Kaufmann and Hollingdale (Vintage, 1967)

Z   *Thus Spoke Zarathustra,* trans. Kaufmann (Viking, 1954)

Bentham, Jeremy. (2007), *An Introduction to the Principles of Morals and Legislation*. New York: Dover Publications.

Bradburn, Norman. (1969), *The Structure of Psychological Well-Being*. New York: Walter de Gruyter.

Clark, Maudemarie and Dudrick, David. (2007), "Nietzsche and Moral Objectivity: The Development of Nietzsche's Metaethics," in *Nietzsche and Morality*, ed. B. Leiter and N. Sinhababu. New York: Oxford University Press.

Csikszentmihalyi, Mihaly. (1988), *Optimal Experience*. New York: Cambridge University Press.

—— (1990), *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

Dancy, Jonathan. (2002), *Practical Reality*. New York: Oxford University Press.

Deci, Edward and Ryan, Richard. (1985), *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Springer.

Dostoyevsky, Fyodor. (2006), *Notes from Underground.* Lawrence, Kansas: Digireads Publishing.

Enoch, David. (2006), "Agency, Schmagency: Why Normativity Won't Come from What is Constitutive of Action." *Philosophical Review* 115: 169–198.

Freud, Sigmund. (1957), *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume 14*, ed. J. Strachey. London: Hogarth Press.

Gass, William. (1971), "The Stylization of Desire." *The New York Review of Books*, 16.3 (February 25).

Kant, Immanuel. (1996), *The Metaphysics of Morals*, trans. M. Gregor. New York: Cambridge University Press.

Katsafanas, Paul. (2008), "Practical Reason and the Structure of Reflective Agency." Harvard University Dissertation.

—— (forthcoming a), "Nietzsche's Philosophical Psychology," in *The Oxford Handbook of Nietzsche*, eds. John Richardson and Ken Gemes. New York: Oxford University Press.

—— (forthcoming b), "The Problem of Normative Authority in Kant, Hegel, and Nietzsche," in a currently untitled volume on Nietzsche and moral philosophy, eds. David Owen and Aaron Ridley. New York: Oxford University Press.

Korsgaard, Christine. (1999), "Self-Constitution in the Ethics of Plato and Kant,," *Journal of Ethics* 3: 1–29.

—— (2009), *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.

Leiter, Brian. (2000), "Nietzsche's Metaethics: Against the Privilege Readings." *European Journal of Philosophy* 8: 277–297.

Mackie, J.L. (1977), *Ethics: Inventing Right and Wrong*. New York: Pelican Books.

McDowell, John. (2001), *Mind, Value, and Reality*. Cambridge: Harvard University Press.

Nehamas, Alexander. (1985), *Nietzsche: Life as Literature*. Cambridge: Harvard University Press.

Poellner, Peter. (2007), "Affect, Value, and Objectivity," in *Nietzsche and Morality*, ed. B. Leiter and N. Sinhababu. New York: Oxford University Press.

Reginster, Bernard. (2006), *The Affirmation of Life*. Cambridge: Harvard University Press.

Richardson, John. (1996), *Nietzsche's System*. New York: Oxford University Press.

—— (2004), *Nietzsche's New Darwinism*. New York: Oxford University Press.

Scanlon, T.M. (2000), *What We Owe to Each Other*. Cambridge: Harvard University Press.

Schopenhauer, Arthur. (1969), *The World as Will and Representation*, trans. E.F.J. Payne. New York: Dover Publications.

Schroeder, Timothy. (2004), *Three Faces of Desire*. New York: Oxford University Press.

Setiya, Kieran. (2007), *Reasons without Rationalism*. Princeton: Princeton University Press.

Smith, Michael. (1994), *The Moral Problem*. Oxford: Blackwell Publishing.

Stalnaker, Robert. (1984), *Inquiry*. Cambridge: MIT Press.

Velleman, J. David. (2000), *The Possibility of Practical Reason*. New York: Oxford University Press.

—— (2009), *How We Get Along*. New York: Cambridge University Press.

Wallace, Jay. (2006), *Normativity and the Will*. New York: Oxford University Press.

Weiss, Bernhard. (1883), *Biblical Theology of the New Testament, trans. D. Eaton*. Edinburgh: T. & T. Clark.

Wilde, Oscar. (1998), *Lady Windermere's Fan*. New York: Dover Publications.

Williams, Bernard. (1981), *Moral Luck*. New York: Cambridge University Press.