

7

The Unified Self

Chapter Six reviewed Nietzsche's arguments against the standard, Kantian model of the will, which treats the will as a faculty enjoying causal independence from the motives and determining action. Although Nietzsche dismisses this Kantian account as trafficking in imaginary faculties and exaggerating the role and importance of conscious thought, he nonetheless gives conscious thought a significant role in the production of action and the shaping of motives. Conscious thought does not make punctual, decisive contributions to the production of action, but does intervene in continuous streams of behavior, altering motives and sometimes transforming this behavior in striking ways. So Nietzsche can draw a distinction between actions in which willing plays a role and those in which it does not; and he can do so without countenancing a substantive faculty called the will.

Yet this model of agency might seem to give rise to a problem. One reason for which philosophers have been eager to locate some faculty of willing is that they think there is a philosophically significant distinction between actions that are preceded or generated by episodes of willing and actions that are not. For example, some philosophers align the willed/non-willed distinction with the free/unfree distinction. Others align it with the morally responsible/not morally responsible distinction. And still others align it with the distinction between genuine action and mere behavior. I will say more about these distinctions; for now, it suffices simply to note that many philosophers have wanted to distinguish actions that are more and less fully our own, and to align the former with the willed actions. However, we might wonder whether the vector model of willing, described in Chapter Six, enables us to draw any such distinctions. After all, actions are the product of a vector of forces that can include drives, affects, and conscious thoughts; when present, these conscious thoughts may play only the smallest of roles in determining the nature of the action. Why, then, should it matter whether and to what extent the conscious thoughts are present? Why not think the question of whether conscious thoughts are present in the etiology of the action is roughly as important as the question of whether a perception of a butterfly is present in the etiology? In short: granted that Nietzsche can distinguish between cases in which conscious willing plays a role in the production of action and cases in which it doesn't, why think this distinction matters?

In this chapter, I'll provide an answer to that question. To put the point simply, I'll argue that in fact it *doesn't* matter whether conscious willing is present in the etiology

of action. Nietzsche doesn't think that there is any philosophically significant distinction between actions whose etiology includes an episode of conscious willing and those whose etiology does not. However, Nietzsche is interested in something that philosophers have (he thinks mistakenly) attempted to capture by speaking of conscious willing: the distinction between *genuine action* and *mere behavior*. Unlike many other philosophers, he does not align this distinction with the willed/unwilled distinction. In other words: genuine actions can be unwilled, and willed actions can be mere behavior.

What then distinguishes genuine actions from mere behaviors? Nietzsche marks the distinction with his concept of *unity*. Genuine actions are those springing from unified agents. And although the unified/disunified distinction does not coincide with the willed/unwilled distinction, I argue that disunity of the relevant sort is possible only for self-conscious creatures—that is, only for creatures who are capable of willing.

These claims about Nietzschean unity will seem surprising to many interpreters. Lately, there has been a profusion of work discussing Nietzsche's views of unity of the self, freedom, and agency.¹ A widespread consensus on three points has emerged: (1) Nietzsche's notion of unity is meant to be an analysis of freedom; (2) unity refers to a relation between the agent's drives or motivational states; and (3) unity obtains when one drive predominates and imposes order on the other drives.

In this chapter, I argue that these three claims are philosophically and textually indefensible. The claims are philosophically indefensible because they fail correctly to characterize certain paradigmatic cases of agency and its absence. The claims are textually indefensible because there are passages indicating that Nietzsche rejects each of them. In contrast to the standard interpretations, I argue that (1') Nietzschean unity is an account of the distinction between genuine actions and mere behaviors, rather than between free and unfree actions; (2') unity refers to a relation between drives and conscious thought; and (3') unity obtains when the agent's attitude toward her own action is stable under the revelation of further information about the action's etiology. I show that Nietzsche develops this account of unity by drawing on Plato's and Schiller's accounts of unified agency; Nietzsche's account incorporates elements of both Plato's and Schiller's accounts, but goes beyond them in emphasizing the roles of non-conscious motivation and the ubiquitous forms of self-ignorance in agency. Accordingly, the account of unified agency that emerges from Nietzsche's works is considerably more sophisticated, and potentially more philosophically fruitful, than has yet been appreciated.

The essay proceeds in five sections. Section 7.1 introduces Nietzsche's criticisms of traditional models of agency, and argues that he aims to replace these accounts with a model of unified agency that overcomes their flaws. Section 7.2 examines Nietzsche's

¹ Many of the relevant essays are contained in Gemes and May (2009). Important earlier works include Mullin (2000), Leiter (2001), Guay (2002), Reginster (2003), Richardson (2009), and Risse (2007).

critique of Platonic unity, and argues against the standard interpretations of Nietzschean unity. Section 7.3 explicates Schiller's notion of unity, and suggests that Nietzsche draws on Schiller's ideas in developing his own conception of unity. Section 7.4 defends this reading by examining a paradigm case of disunity, the *Genealogy's* ascetic priest, and drawing from it an account of what unity must be. Section 7.5 examines the broader significance of Nietzsche's notion of unity.

7.1 The Philosophical Role of the Concept of Unity

7.1.1 Nietzsche's Critiques

Consider a few descriptions of deliberative agency in the contemporary literature. Christine Korsgaard describes the Kantian model of agency as follows: "when you deliberate, it is as if there were something over and above all of your desires, something which is you, and which chooses which desire to act on" (Korsgaard 1996, 100). Other writers put the point differently, but agree with the underlying idea. For example, Michael Bratman writes:

When a person acts because of what she desires, or intends, or the like, we sometimes do not want to say simply that the pro-attitude leads to the action. In some cases, we suppose, further, that the agent is the source of, determines, directs, governs the action and is not merely the locus of a series of happenings, of causal pushes and pulls. (Bratman 2007, 91).

Bratman here claims that we need to distinguish the operations of the agent from the operations of the agent's attitudes. These philosophers invoke the idea of an agent or a self, who serves as something more than a mere container for the various desires and affects that cause acts. In cases of genuine action, the agent is somehow the "source of" the act. Yet Nietzsche denies that there is any self over and above the drives. The self, Nietzsche tells us, is just a "relation" or "social structure of drives and affects" (BGE 12). Accordingly, it is difficult to see how we could draw a distinction between acts caused by agents and acts caused by drives and affects.

Although Nietzsche is deeply skeptical both of the idea that reflective choice is independent from drives, and of the notion of selfhood typically employed in accounts of agency, it would be a mistake to conclude that he rejects the distinction between genuine agency and mere undergoing.² On the contrary, Nietzsche is explicit about his reliance on such a distinction. Not only does Nietzsche tell us that activity [*Aktivität*] is one of his foundational concepts [*Grundbegriffe*] (GM II.12), he also repeatedly relies upon a distinction between genuine actions and their degenerate relatives. Thus, he praises the "sovereign" or "autonomous" individual, who is distinguished by the fact that he "has his own independent, protracted will" (GM II.2). While the acts of

² For the moment, I want to remain vague on what exactly "genuine agency" is. I will explicate this notion in Section 7.1.2.

non-sovereign individuals are simply determined by whatever impulse happens to arise, the acts of sovereign individuals are controlled by the agent herself. For the sovereign individual is “strong enough to maintain [her commitments] even in the face of accidents, even ‘in the face of fate.’” By contrast, the non-sovereign individual is “short-willed and unreliable,” he “breaks his word even at the moment he utters it.” For the non-sovereign individual is incapable of holding himself to a course of action in the face of accidents and temptations. Unable to regulate his own behavior, the non-sovereign individual will only fulfill his projects and goals if, through sheer luck, he encounters no temptations.

As we saw in Chapter Six, Nietzsche elsewhere develops these ideas, claiming that some agents have the power “not to react at once to a stimulus, but to gain control of all the inhibiting, excluding instincts . . . the essential feature is precisely not to ‘will,’ to be able to suspend decision. All unspirituality, all vulgar commonness, depend on an inability to resist a stimulus: one must react, one follows every impulse” (TI VIII.6). In the same work, Nietzsche defines weakness as the “inability not to respond to a stimulus” (TI V.2). The weak individual’s actions are determined by whatever impulse or stimulus happens to arise; she possesses no capacity to direct her own behavior. By contrast, the strong individual is able to check her impulses and resist stimuli.³

In these passages, Nietzsche claims that some individuals have the capacity to control their behavior. Moreover, these characterizations of agency seem to invoke the very images that Nietzsche elsewhere rejects: they suggest a self that stands apart from the drives and impulses, and exerts a controlling influence over them.

Nietzsche’s characterizations of strong, sovereign agents controlling their actions suggest that there is something correct in our ordinary distinction between acts actively produced by the agent and acts in which the agent is a mere vessel for forces within. At the same time, Nietzsche critiques the way in which agency is normally understood. So he must have some alternative way of characterizing agency.

Is there a way of drawing a distinction between genuine agency and its lesser relatives, without relying on the idea of a self independent from the drives, and without denying that drives exert a pervasive influence on choice? I will argue that there is. Contemporary philosophers who attempt to distinguish genuine actions from their lesser relatives sometimes appeal to a distinction between being a genuine agent and

³ There are a number of similar passages. Consider a few passages from Nietzsche’s notebooks. In KSA 12:9[178], Nietzsche condemns nineteenth-century thinkers for being “deeply convinced of the rule of cravings [*Begierde*]. (Schopenhauer spoke of ‘will’; but nothing is more characteristic of his philosophy than the absence of all genuine willing).” KSA 13:11[353] speaks of great individuals controlling their affects: “Greatness of character does not consist in not possessing these affects—on the contrary, one possesses them to the most terrifying [*furchtbarsten*] degree—but in having them under control [*aber dass man sie am Zügel führt*].” KSA 12:9[139]/WLN 163 makes a similar point: “In summa: to master the passions, *not* to weaken or exterminate them!—The greater the mastering power of a will, the more freedom may the passions be allowed. The ‘great man’ is great owing to the free play and scope of his desires and to the yet greater power than knows how to press these magnificent monsters into service.” Cf. GM II.3.

being a mere locus of forces. Nietzsche does not accept the distinction between being an agent and being a locus of forces; indeed, the vector model of willing described in Chapter Six treats the agent as a locus of forces. However, Nietzsche does distinguish between *different kinds* of loci of forces.

In particular, Nietzsche tells us that some agents are disunified loci of forces, whereas other agents are unified. Thus, Nietzsche argues that agents are typically multiple and fragmented. He notes that “human beings have in their bodies the heritage of multiple origins, that is, opposite, and often not merely opposite, drives and value standards that fight each other and rarely permit each other any rest” (BGE 200). As a result, “our drives now run back everywhere; we ourselves are a kind of chaos” (BGE 224). Thus, “the belief which regards the soul as . . . a monad, as an *atomon*: this belief ought to be expelled from science! . . . But the way is open for new versions and refinements of the soul-hypothesis; and such conceptions as ‘mortal soul,’ and ‘soul as subject-multiplicity’ [*Subjekts-Vielheit*], and ‘soul as social structure of the drives and affects,’ want henceforth to have citizens’ rights in science” (BGE 12). These sections have a common theme: they deny that there is a *unitary* self, and assert that the self should be understood as a relation, complex, or social structure of drives.⁴

However, Nietzsche makes it clear that this is a contingent state. While most individuals are disunified, attaining a unified self is possible: Nietzsche claims that modern individuals can be made “whole” or unified (TI IX.41), and he cites Goethe as an example of someone who “disciplined himself to wholeness” (TI IX.49). Thus, although the self is typically disunified, it is possible to attain a unified self.

In sum, while Nietzsche rejects the distinction between being an agent and being a locus of forces, he introduces a distinction between disunified and unified loci of forces. I will argue that Nietzsche’s account of unity is an account of genuine agency: what it is for an agent to play an active role in producing her action is for the agent to be unified in acting.

7.1.2 *Three Distinctions among Doings*

Before proceeding, though, it will be necessary to clarify what exactly Nietzsche’s account of unified agents is meant to be an account of. Philosophers of action typically draw at least three distinctions between the movements issuing from an agent: mere behavior, action, and autonomous (or free) action. On most accounts of agency, autonomous action and action can come apart: when I give the mugger my wallet, or when I act unreflectively out of mere custom or habit, I am *acting*, but I may not be acting *freely*.⁵ So we can distinguish, among the set of actions, between the free and unfree ones.

⁴ For a helpful discussion of these points, see Gemes (2001).

⁵ Here I am assuming that coercion and habitual action are examples of unfreedom. Of course, not all accounts of freedom will classify these as unfree actions. Readers who prefer an alternative account of freedom can substitute examples of their own.

However, we can also distinguish, among the set of movements that issue from the agent, between *mere behaviors* and *genuine actions*. Consider the distinction between movements such as sneezing, coughing, falling asleep, and blinking, on the one hand, and reading, conducting conversations, getting married, and deciding to go to Bermuda, on the other. Each of these events counts as something that a person does, in a sense, but there seem to be important differences between, say, sneezing and getting married. The former is a reflex, something that happens to me, something that is not entirely under my control; the latter is a product of choice, something that I do, something that is to some extent under my control. We can mark this distinction by calling the sneeze and its ilk *mere behaviors*, and the marriage and its ilk *actions*.

Many philosophers believe that the class of mere behaviors includes not only reflex behaviors of the sort mentioned above, but also a variety of behaviors that seem to be brought about independently of the agent's reflective thought or deliberation. For example, David Velleman claims that any case in which an agent is *ignorant* of her action, or in which she discovers what she is doing *only by observing herself doing it*, counts as a mere behavior (Velleman 2000).⁶ Harry Frankfurt argues that any case in which the agent is not *identified with* his motives is a mere behavior (Frankfurt 2004). The idea, here, is that full-fledged actions require the agent to play an active role in the production of her own activity; anything less counts as mere behavior.

Nietzsche seems to have something like this distinction in mind when he writes, "Nothing is rarer than a *personal* action. A class, a rank, a race, an environment, an accident—everything expresses itself sooner in a work or deed, than a 'person' [*Alles drückt sich eher noch in einem Werke oder Thun aus, als eine „Person“*]" (KSA 12:10[59]). Here, Nietzsche claims that what appears to a case of a person actively bringing about an action is better described as some force acting through the person. Or, to put the point in contemporary terminology, what looks like action is really mere behavior.⁷

In sum, we have three distinctions: mere behaviors, actions, and free actions. So we should ask which of these distinctions Nietzsche's concepts of unity and disunity are meant to mark. Does the disunified/unified distinction correspond to the mere behavior/action distinction, or to the unfree action/free action distinction?⁸

Commentators on Nietzsche have assumed that his remarks on unity are meant to distinguish free actions from unfree actions: an act is free if the agent is unified, otherwise the act is unfree.⁹ However, I will argue that this is a mistake. Nietzsche's remarks

⁶ Velleman actually uses the term "mere activity," rather than "mere behavior," to describe these cases.

⁷ D 38 makes a similar distinction.

⁸ Of course, a third possibility is that the remarks on unity and disunity do not correspond to any of these categories. The evidence in the following sections counts against this reading.

⁹ The following works endorse this view: Nehamas (1985), Reginster (2003), Risse (2007), and Thiele (1990). Gemes (2009a) also reads Nietzsche as identifying freedom with unity. However, Gemes distinguishes "agency free will" from "deserts free will." Whereas deserts free will focuses on the questions of desert, punishment, and responsibility, agency free will focuses on the question of "what constitutes an action as opposed to a mere doing." Thus, what Gemes is calling agency free will is close to what I am calling the distinction between action and mere behavior. Gemes claims that Nietzsche is interested in agency free will, not deserts free will. My interpretation is in agreement with most of these points. However,

on unity are meant to distinguish actions from mere behaviors: what makes something an action, as opposed to a mere behavior, is that the agent is unified. On this reading, there is a further question concerning whether all unified actions are free actions: given that freedom is more demanding than unity, an agent could be unified without being free.

I think there is unambiguous textual evidence establishing that this is Nietzsche's view: unity is a necessary, but not a sufficient, condition for freedom.¹⁰ To see this, notice that Nietzsche's exemplars of unity are, in general, the members of certain castes in ancient societies. Throughout his works, Nietzsche emphasizes that these individuals were free from the kinds of inner conflict that plague modern human beings: "they were more *whole* human beings" (BGE 257). Nietzsche claims that certain forms of social organization and certain systems of morality—the ones that were dominant in Homeric Greece, for example—are conducive to the production of unified individuals; other forms of social organization and moral systems, such as the ones dominant today, are conducive to the production of fragmented, conflicted persons. Accordingly, the Greeks of the Homeric era were, in general, unified individuals, whereas moderns are, in general, disunified individuals.¹¹

However, when Nietzsche is discussing *freedom*, he often singles out traits that seem entirely absent in the Greek nobility. The discussions of freedom focus on "evaluating on one's own," "revaluating," creating new values, questioning traditional values, and actively inquiring into the history and the effects of values. These are not the traits that spring to mind when we consider an Achilles or an Agamemnon. The Homeric nobles are paradigms of physical strength, health, self-assertion, and self-certainty. But they are clearly *not* examples of individuals struggling to gain independence from traditional values, or to win truth or self-understanding. Indeed, Nietzsche explicitly states that the critical stance involved in the pursuit of truth and the questioning of tradition is a distinctively modern achievement. As he puts it, "the ability to contradict, the attainment of a good conscience when one feels hostile to what is accustomed, traditional, and hallowed—that is still more excellent and constitutes what is really great, new, and amazing in our culture" (GS 297). Accordingly, while figures such as Achilles are unified, they do not seem to be *free*.¹²

Gemes' approach collapses the threefold distinction between mere behavior, action, and free action into a twofold distinction between non-action and action. I think Nietzsche does have a threefold distinction, as I will argue below.

¹⁰ John Richardson and Simon May accept versions of this view. Richardson argues that freedom requires both unity and genealogical insight, whereas May argues that freedom requires both unity and the overcoming of nihilism. See Richardson (2009) and May (2009).

¹¹ For discussions of the way in which ancient societies tended to produce unified individuals, whereas modern societies tend to produce disunified individuals, see BGE 257, the whole of GM, and HC.

¹² On rare occasions, and mostly in the very early texts, Nietzsche does claim that certain ancients strove for self-determination. For example, in *Philosophy in the Tragic Age of the Greeks*, he tells us that "the freedom of the individual was in [Xenophanes] at its zenith," because he engaged in a "daring disapproval of the existing customs and valuations." That is, Xenophanes engaged in an attempted revaluation of values.

Thus, Nietzsche's exemplars of unity do not seem to possess the traits that are characteristic of free individuals. This suggests that unity and freedom are distinct. And in fact this point becomes clear when we consider one trait in particular: the free individual is said to be *liberated from* or *independent of* morality. As Nietzsche puts it in the *Genealogy*:

The ripest fruit is the *sovereign individual*, like only to himself, liberated again from the morality of custom [*Sittlichkeit der Sitte*], autonomous and supramoral [*übersittliche*] (for 'autonomous' and 'moral' are mutually exclusive), in short, the man who has his own independent, protracted will. (GM II.2)

The sovereign individual is *liberated from the morality of custom*; he is autonomous, and therefore, Nietzsche claims, "supramoral." But Nietzsche believes that the ancient Greeks are precisely those who *embraced* the morality of custom. A few lines before the passage quoted above, Nietzsche refers the reader to his discussions of "the morality of custom" in *Daybreak*. Turning to the relevant passages in *Daybreak*, we learn that "all the communities of mankind . . . up to the present day" have lived under the "morality of custom" (D 14). Moreover, Nietzsche notes that Socrates was one of the first to attempt to step beyond the morality of custom (D 9). This implies that, on Nietzsche's view, the pre-Socratic Greek nobles lived under the morality of custom. While they were unified individuals, they could not have been free individuals.

Nietzsche must, then, distinguish between unity and freedom. One can be unified without being free: unity is compatible with uncritical, unreflective government by certain forms of morality, whereas freedom is not.¹³ So unity is at best a necessary condition for freedom. For this reason, I take it that Nietzsche's distinction between unity and disunity is meant to capture the distinction between mere behavior and action, rather than the distinction between action and free action. (I will return to the notion of freedom in Chapter Eight.) We can see these concerns arising in the following passages:

"I have no idea what I am *doing*! I have no idea what I *ought to do*!"—You are right, but be sure of this: *you are being done!* [*du wirst gethan!*] at every moment! Mankind has in all ages confused the active and the passive: it is their everlasting grammatical blunder. (D 120)

Nothing is rarer than a *personal* action. A class, a rank, a race, an environment, an accident—everything expresses itself sooner in a work or deed, than a "person." (KSA 12:10[59])

In both passages, Nietzsche seeks a distinction between the agent's playing an active role in the determination of her action and the agent's being a passive vessel for external

However, Nietzsche's considered view seems to be that self-determination is a distinctively modern achievement. See HH Preface 3, HH 225; D 9; GS 347; GM II.2; BGE 44, 211, 227; A 54; EH IV.1.

¹³ Nietzsche would not claim that, for any evaluative system, one can embrace that evaluative system and be unified. Rather, he claims that there are certain evaluative systems the adoption of which is compatible with unity. For example, he suggests that the evaluative systems of certain classical societies were compatible with unity, whereas Judeo-Christian evaluative systems are not compatible with unity.

forces. And again, I will suggest that he wants to mark this distinction by appealing to a certain form of unity.

7.2 Nietzsche and Plato on the Unified Agent

At this point, we know what philosophical role the concept of unity is meant to play: it marks the distinction between genuine action and mere behavior. However, we still need to determine what unity *is*.

The claim that the self is initially multiple, and that unification is an achievement, has a distinguished philosophical pedigree: we find it already in Plato. It will be useful to begin by contrasting the Platonic model with the Nietzschean model, for two reasons. First, Nietzsche self-consciously opposes his model to the Platonic model. Second, I will argue that many commentators have failed to recognize crucial respects in which Nietzschean unity differs from Platonic unity.

7.2.1 *Nietzsche's Departures from Plato*

To begin, we will ask two questions about the Platonic and Nietzschean models. First, what are the parts into which the self is divided? Second, what kind of relation among these parts is required in order for unity to be achieved?

In the *Republic*, Plato claims that the soul has three parts: Reason, Appetite, and Spirit (*Republic* 580d–581d). There is scholarly controversy regarding the nature of these parts, for Plato sometimes treats them as faculties or independent parts of the self, sometimes as values, and sometimes as the objects of strong dispositions. Thus, in a recent discussion of Plato's tripartite soul, Ferrari writes that "in Book 4 they [the parts] seemed most like faculties; now [in books 8 and 9] they seem more like drives. The desiring element is specified as the drive toward material satisfaction; spirit as the drive to win and to amount to something; calculation as the drive to discover truth" (Ferrari 2007, 165). The term "drive" is being used loosely, here, to mean something like an overriding motive or disposition; it does not have the Nietzschean sense of a disposition that admits an aim/object distinction. According to Ferrari, then, Plato is claiming that there are three dominant sources of motivation in the agent, and these motivational sources can be configured in different ways.

We needn't resolve these interpretive controversies over the nature of the Platonic parts; it won't matter, for our purposes, whether Plato envisions the parts as independent faculties or merely as strong sources of motivation. The important point, for our purposes, is that certain relations among these parts render the agent disunified, whereas others render the agent unified. In particular, the agent is unified when Reason exerts a dominant influence over the agent's action, and disunified when Appetite or Spirit reigns. As Plato puts it, it is "appropriate for the rational part to rule, since it is really wise and exercises foresight on behalf of the whole soul" (*Republic* 441e; cf. 442d).

How we understand the claim that Reason should rule will depend on what we take Reason to be. If we interpret Reason as a faculty, we would have something like the

Kantian picture of action: Plato would be claiming that reflective, rational thought should set our ends, whereas affective elements of the self (Appetite and Spirit) should be subordinated. If, by contrast, we interpret Reason merely as a motivational source, Plato would be claiming that a motive toward discovering truth should dominate the motives toward material satisfaction and toward competition.

Much more could be said about Plato's view, but for our purposes this spare characterization will suffice. We can view Plato's account as consisting of two claims:

- (1) Platonic Parts: The self is divided into three parts: Reason, Appetite, and Spirit.
- (2) Platonic Relation: The self is unified when one of these parts, Reason, dominates the other parts.

Does Nietzsche agree with Plato? Well, one difference is immediately obvious. Whereas Plato thinks the soul has three parts, Nietzsche is much more profligate: throughout his corpus he names over one hundred distinct drives. Some commentators believe that this is the primary difference between the Platonic soul and the Nietzschean self: the Nietzschean self has *more* parts. For example, Thiele writes, "Platonic opposition between reason and passion is fractured into the opposition between multiple passions, each with its own capacity for reason and will to dominate" (Thiele 1990, 56). Clark and Dudrick echo this point, writing that their "claim is not that Nietzsche *accepts* Plato's theory of the soul but that he takes it to be a hypothesis, which his own psychology is an attempt to refine" (Clark and Dudrick 2012, 163). They go on to claim that "Nietzsche takes *who the person is* to be constituted by the political order of her drives" (2012, 177).

On this reading, Nietzsche basically agrees with Plato, differing only on the *number* of parts. Thus, on Thiele's interpretation, Nietzsche modifies claim (1) in the following way:

- (1') Nietzschean Parts: The self is divided into more than three parts. These parts are drives.

Claim (1') is a common interpretation of Nietzsche, shared by Mathias Risse, Ken Gemes, Christopher Janaway, and others.¹⁴ Sections 7.3 and 7.4 will argue that (1') is false: the parts into which Nietzsche divides the self include more than just drives. While the drives jointly compose one part of the self, there is another part: the agent's self-conscious thought. For now, though, let us proceed to claim (2).

The texts make it clear that Nietzsche disagrees with (2). Throughout his works, Nietzsche inveighs against Plato's claim that Reason should dominate the other parts: Plato "turn[ed] *reason* into a tyrant" (TI II.10).¹⁵ Although Nietzsche repeatedly claims that Plato is mistaken in endorsing Reason's dominance over the other parts, Nietzsche

¹⁴ See Thiele (1990), Risse (2007), Gemes (2009a), Clark and Dudrick (2012), and Janaway (2014).

¹⁵ For further remarks to this effect, see TI II.1–12, TI V.1, BGE 191, and KSA 12:9[166].

is not as explicit as one would like about *why* this is a mistake. Let me pause for a moment over this point.

Plato is envisioning a mode of life in which a drive toward truth as such dominates other sources of motivation. As Janaway puts it, Plato's claim is that:

There is a basic element in us with its own passionate interest in knowledge, an element that responds affectively to seeking and uncovering truth... It is a persisting disposition to value knowledge of the truth as an end above all other ends. It is what we may call a pure drive toward truth. (Janaway 2014, 16).

The best life, for Plato, is the one in which this pure drive toward truth subordinates the agent's other drives. However, from Nietzsche's perspective there are two deep problems with this view.

First, Nietzsche argues that Reason cannot be disentangled from Appetite. As the previous chapters discussed, Nietzsche argues that the agent's rational faculties are pervasively influenced by drives and affects. Nietzsche therefore concludes that the dominance of Reason, as Plato understands it, is impossible: any putatively "pure" drive toward truth will contain a mixture of other motivations (see, for example, GM III's emphasis on the way in which scientists and philosophers, who appear to be driven solely by a quest for truth, are motivated by ascetic tendencies). This fact renders claims about Reason's dominance problematic: if we cannot disentangle Reason from various appetites, then the claim that Reason should dominate Appetite seems untenable.¹⁶

Second, Plato claims that there is only *one* way to achieve unity: Reason must predominate. Nietzsche suggests that there are *many* different ways to attain unity. There is no one drive that must dominate, in order for the soul to be unified. Rather, commentators often interpret Nietzsche as arguing that unity obtains when one drive—*any* drive—exerts a dominant influence. For example, Gemes claims that "it is when a strong will [i.e., drive] takes command, orders and organizes lesser drives" that a person manifests genuine agency (Gemes 2009a, 42). Richardson suggests that Nietzsche associates unity with the emergence of "a single dominant drive, or perhaps a ruling 'committee' of drives" (Richardson 2009, 134–5). Thiele claims that unity is achieved "through the harnessing of violent and varied passions, and their placement under the rule of a predominant drive" (Thiele 1990, 63). And Clark and Dudrick argue that unity is attained when subordinate drives recognize a predominant drive as having (as they put it) "political authority" (Clark and Dudrick 2012, 183; cf. 175–95).

Thus, a number of commentators assume that while Nietzsche disagrees with the details of Plato's account, he accepts Plato's basic claim that the soul is unified when one part dominates and imposes order on the other parts. Suppose we break Plato's claim into two parts:

¹⁶ Moreover, Janaway offers an interesting argument that even if there were a drive toward knowledge, Nietzsche would not treat its predominance as leading to the best life. See Janaway (2014, 18–19), as well as BGE 6 and GM III.28.

- (2a) Platonic Relation: The self is unified when one part dominates, and
- (2b) This part must be Reason.

On the standard reading of Nietzschean unity, Nietzsche accepts (2a) and rejects (2b).

However, I think this is a mistake. Nietzsche's departure from Plato is more radical, for he also rejects (2a). He denies the idea that unity is achieved via one part's dominance over the other parts. While Nietzsche does think that dominance of one part *frequently causes* unity, he denies that dominance is the *same thing as* unity. Below, I argue for this point. I will first consider textual evidence in favor of the standard reading of Nietzschean unity, which claims that unity obtains when any drive exerts a dominant influence on the other drives. I will then offer some philosophical and textual reasons for rejecting this interpretation.

7.2.2 Textual Evidence for the Claim That Unity Is the Predominance of One Part

Consider two passages from Nietzsche's *Nachlass*:

The multiplicity and disaggregation of drives, the lack of any systematic organization among them, results in a "weak will"; their coordination under the dominance [*Vorherrschaft*] of a single one results in a "strong will"; in the first case it is the oscillation and lack of gravity; in the latter, the precision and clarity of direction. (KSA 13:14[219]/WLN 266)

The antagonism of the passions: two, three, a multiplicity of "souls in one breast": very unhealthy, inner ruin, disintegration, revealing and increasing an inner conflict and anarchism—unless one passion at last becomes master. *Return of health*—. (KSA 13:14[157]/ WLN 262)

In the first passage, Nietzsche tells us that the will is strong when one drive predominates and coordinates the other drives. If we assume that Nietzsche's talk of "strong" wills is meant to refer to unified wills, then Nietzsche seems to be offering a straightforward analysis of unity: an agent is unified if one of his drives coordinates the other drives. The second passage makes a similar point: an agent is unhealthy and experiences inner conflict when there is no drive that has become master (i.e., no drive that is dominant).

This seems plausible. The predominance of one drive seems to provide opportunities for locating the agent in the production of action. In particular:

- (a) We could identify the agent's acts with the acts caused by the dominant drive.
- (b) Alternatively, if the agent's will were identified with the whole set of drives, then the agent's will would be unified.

By contrast, consider an individual whose drives exhibit anarchy: the drives are inconsistent, at odds with one another, and no one drive is predominant. Such an individual would lack both of the conditions described above:

- (a') There is no *one* drive whose operations could be regarded as expressive of the agent. For no drive enjoys predominance over the others.

- (b') If the agent's will were identified with the whole set of drives, then the agent's will would be disunified, a chaotic mix of warring fragments.

So we can see why it is tempting to think that predominance of one drive constitutes unity.

There is, however, a complication. Nietzsche frequently praises individuals who harbor diverse, inconsistent, conflicting drives:

Man has, in contrast to the animals, cultivated an abundance of *opposed* [gegensätzlicher] drives and impulses within himself: thanks to this synthesis, he is master of the earth... The highest man would have the greatest multiplicity of drives, in the relatively greatest strength that can be endured. Indeed, where the plant 'man' shows himself strongest one finds drives that conflict powerfully (e.g., in Shakespeare) but are restrained [*gebändigt*]. (KSA 11:27[59])

A philosopher—if today there could be philosophers—would be compelled to find the greatness of man, the concept of 'greatness', precisely in his range and multiplicity, in his wholeness in manifoldness... Precisely this shall be called *greatness*: being capable of being as manifold as whole, as ample as full [*ebenso vielfach als ganz, ebenso weit als voll sein können*]. (BGE 212)

On the face of things, these passages count against the idea that unity consists in dominance of one drive. Nietzsche's exemplars of unified selfhood—above he mentions Shakespeare, and elsewhere Goethe, Napoleon, and Nietzsche himself are the paradigms—are praised precisely because they have diverse, powerful, and inconsistent drives, but are *in some other sense unified*.

However, proponents of the predominance model claim that we can account for this point by distinguishing different forms of predominance by one drive. Chief among these are tyranny and mastery.¹⁷ Tyranny consists of one drive's achieving a predominant status by suppressing or extirpating other drives (HH I.228, GS 347). In other words, drive A tyrannizes drives B and C when A becomes stronger than B and C by weakening or eliminating B and C. Nietzsche offers asceticism as a paradigmatic form of tyranny. Mastery, by contrast, consists of one drive's being predominant, but still allowing other drives expression. In other words, drive A masters drives B and C when A becomes stronger than B and C, and modulates the expression of B and C, yet does not weaken or eliminate B and C. An example might be a dominant drive toward intellectual activity modulating the expression of, say, the aggressive drive and the sex drive; the intellectual drive might master these drives in the sense that the agent allows the latter drives expression only when doing so does not interfere with the expression of the intellectual drive. The idea, then, would be that Shakespeare, Goethe, and Nietzsche's other exemplars harbor a drive that predominates through mastery.¹⁸

¹⁷ Again, the connection to Plato should be clear. Plato distinguished five different types of unity and disunity: aristocracy, timocracy, oligarchy, democracy, and tyranny.

¹⁸ Reginster (2003) discusses the notions of anarchy, tyranny, and mastery in detail. See also Gemes (2009b). Gemes uses the terms "sublimation" and "repression" to pick out conditions similar to those that Reginster labels "mastery" and "tyranny." Employing these notions, Gemes provides a very helpful discussion of the possible configurations of drives.

Thus, we seem to have an account of unity. The self is unified when one drive predominates, and exerts a coordinating influence on the other drives. Conflict amongst drives does not have to be *eliminated*; it just has to be *managed*.

7.2.3 *Textual and Philosophical Reasons for Rejecting the Reading of Unity as Predominance*

The predominance model seems to provide a tidy answer to the question of what Nietzsche's account of unity is. However, I think that upon examination this account becomes far less plausible, for two reasons.

First, there is an obvious philosophical problem with the view that unity is dominance by one part of the self: it assumes that the dominant part of the self has some special claim to being expressive of the self. But this assumption is unwarranted, for we often distinguish actions produced by the agent from acts caused by the agent's dominant motive. For example, imagine an alcoholic who ardently craves a drink, but judges that he should resist. Suppose the craving eventually overpowers the agent's resistance. Here, the craving for alcohol is the strongest motivational force, but it would be perverse to say that when the agent acts on that craving, he manifests agential control. On the contrary, the voice of the agent seems to reside in the weak, overpowered element of resistance. It is for good reason, then, that we distinguish acts that are expressive of the agent from acts that are expressive of the strongest motivational force.

A proponent of the unity-as-predominance model might respond to this objection by claiming that the alcoholic urges should be regarded as operating through tyranny rather than mastery. That is, if unity requires predominance in the form of mastery, and if alcoholics exhibit predominance only in the form of tyranny, then alcoholics would not serve as counterexamples to the unity-as-predominance model.

However, this response on behalf of the predominance model does not seem promising. In order for the response to succeed, one would have to show that there are *no* cases in which alcoholism operates via mastery. This is implausible. Recall that mastery simply requires that one drive modulates, but does not weaken or extirpate, other drives. A number of alcoholics seem to fit this description. For example, "high-functioning" alcoholics are defined as those who maintain stable and successful lives, often for many decades, despite an addiction to alcohol. Many of these agents have rich arrays of passions and drives that are subordinated to, but not weakened or extirpated by, their craving for alcohol.¹⁹ Thus, the alcoholic urges should count as mastering the other

¹⁹ Ernest Hemingway is sometimes cited as an example of a high-functioning alcoholic. Hemingway reportedly drank a quart of liquor per day for most of his adult life. Sometimes he drank even more: over one sixth-month period in the early sixties, he seems to have consumed eighteen bottles of liquor and 660 bottles of wine. Aside from the sheer amount that he drank, the signs of alcoholism were clear: he was unable to stop drinking when his doctors advised him to do so, and he seemed highly dependent on alcohol—witness his claim that "You wake up in the night and things are unbearable and you take a drink and make them bearable" (quoted in Meyers 1999, 539). This intense drinking seems to have begun in the 1920s. The physiological effects started to appear in the 1950s, and progressively worsened until Hemingway's suicide in 1961—a reporter who interviewed Hemingway in 1958 noted that Hemingway's

drives. If the proponent of the predominance model insists that these high-functioning alcoholics, despite their manifold accomplishments and varied activities, are tyrannized by their drive for alcohol, then one starts to lose any grip on what the distinction between tyranny and mastery is supposed to be.

In further support of this point, notice that high-functioning alcoholism usually does not last for the individual's whole life. In typical cases, the alcoholic's condition lasts for several years and sometimes even several decades, but eventually deteriorates. The alcoholic becomes incapable of maintaining his usual routines: he loses his job, his other passions, his friends, and so on. After this collapse in functioning, the craving for alcohol operates in a very different way: the craving extirpates or severely weakens competing drives, and becomes all-consuming. Given the definitions of mastery and tyranny, the correct characterization seems to be this: prior to the collapse in functioning, the alcoholic craving operates via mastery; after the collapse, it operates via tyranny.

I conclude that certain alcoholics do indeed serve as counterexamples to the unity-as-predominance model. Alcoholics who are mastered by their alcoholic urges would count as unified according to the predominance model, and yet it seems perverse to claim that when such an alcoholic succumbs to his addiction he is manifesting agential control.

So the first problem with the predominance model is that we often distinguish acts caused by the agent's strongest motive from acts produced by the agent herself. But there is also a second problem with the predominance model: there is textual evidence that Nietzsche dissociates unity and dominance. After all, he derides those of us who, like the alcoholic, become "as a whole the victim of some detail of us [*als Ganzes das Opfer irgend einer Einzelheit an uns werden*]" (BGE 41). More decisively, one of Nietzsche's paradigms of dominance by one drive is *also* a paradigm of disunity. Nietzsche claims that Richard Wagner's personality is controlled by one drive:

The dramatic element in Wagner's development is quite unmistakable from the moment when his ruling passion became aware of itself and took his nature in its charge: from that time on there was an end to fumbling, straying, to the proliferation of secondary shoots, and within the most convoluted courses and often daring trajectories assumed by his artistic plans there rules a single inner law, a will by which they can be explained. (UM III.2)

Nietzsche here claims that Wagner's rich and diverse personality, with manifold interests and passions, was dominated by one ruling passion or drive. Yet Nietzsche treats Wagner as a paradigm of *disunity*: he is "the *most instructive case*" for a philosopher

"liver was bad. You could see the bulge of it stand out from his body like a long, fat leech" (Meyers 1999, 539). Nevertheless, for many decades Hemingway maintained a rich, varied, and productive life. By any plausible standard, we must conclude that his alcoholism allowed his other drives expression—after all, he maintained an active social life, he flourished as a writer, he had diverse intellectual interests, he traveled extensively, and so on. Hemingway thus seems to be a paradigmatic case of an individual whose alcoholic cravings master, rather than tyrannize, the other drives. Nonetheless, Hemingway seems passive in the face of his alcoholism.

interested in a “*diagnosis of the modern soul*” (CW, Epilogue; italics in original). As we saw above, Nietzsche defines the “modern soul” as the disunified soul. If Wagner is supposed to be an exemplar both of one drive’s dominance *and* of disunity, then unity cannot be identical with one drive’s dominance.²⁰

One might object that the passage claiming that Wagner is controlled by one drive is from a very early work, whereas the passage claiming that he is disunified is from a late work. However, this objection fails: the claim that Wagner is dominated by one drive is present not only in early works such as UM, but also in some of Nietzsche’s very last works. For example, in CW Nietzsche repeatedly states or implies that Wagner is dominated by one drive. Nietzsche writes that “one cannot figure out Wagner until one figures out *his dominant drive*” (CW 8; cf. CW 11). He explains:

One does not understand a thing about Wagner as long as one finds in him merely an arbitrary play of nature, a whim, an accident. He was no ‘fragmentary’, ‘hapless’, or ‘contradictory’ genius, as people have said... If anything in Wagner is interesting it is the logic with which a physiological defect makes move upon move and takes step upon step in practice and procedure, as innovation in principles, as crisis in taste. (CW 7)

Nietzsche’s suggestion, then, is that some drive or “physiological defect” dominated Wagner and imposed an overall order on his life. In short, throughout his works Nietzsche treats Wagner as both controlled by one drive and disunified.

Thus, there are compelling philosophical and textual reasons for distinguishing unity and the predominance of one drive. So we need a new account of unity.²¹

7.2.4 Clark and Dudrick on Unity as Political Order

Clark and Dudrick (2012) offer a different solution. They argue that the self comprises drives that are normatively or politically ordered, rather than causally ordered (Clark and Dudrick 2012, 139). They reject the idea that “the causal order” of the drives explains the type of person one is, why the person holds her values, and why she holds her theoretical views (2012, 149). Instead, they suggest that Nietzsche “takes the soul to be ‘the political order of the drives and affects’... The political order is not just a causal order, but a normative one” (2012, 150). Although they acknowledge that Nietzsche “never comes right out and says that it [the ordering of drives] is... a normative rather than a merely causal order” (2012, 155), they claim that we can make sense of his claims only by interpreting him in this way.

²⁰ See also GM III.4, where Nietzsche claims that Wagner experienced “a deep, thorough, and even frightful identification with and descent into medieval soul-conflicts.”

²¹ I do not deny that Nietzsche is interested in the psychic conditions of predominance. My claim is simply that Nietzsche does not identify the condition of predominance with the unity that is necessary in order for the agent to play an active role in producing the action. The account of mastery is nothing more than what Nietzsche explicitly says: it is an account of “strong” or “healthy” wills (KSA 13:15[157]/WLN 261, KSA 13:14[219]/WLN 266). But a will can be strong, in this sense, without *the agent* playing any role in the production of the action.

This interpretation relies on several contentious claims: that Nietzsche recognizes a normative/causal distinction; that drives are the kinds of entities that could stand in normative relations to one another; and that, if they could, this would resolve the problems discussed above. I will discuss these claims in turn.

Consider, then, the claim that Nietzsche would distinguish causal and normative orders. Clark and Dudrick admit that there is not a single text in which Nietzsche explicitly endorses this distinction. Their evidence that he accepts the distinction is therefore exceedingly thin: they note that Nietzsche spent a great deal of time studying Afrikan Spir's text *Denken und Wirklichkeit*, which does explicitly endorse a normative/causal distinction. Moreover, they claim that making this distinction enables us to understand why Nietzsche criticizes mechanistic science. However, both of these arguments are unconvincing.

First, Spir is, as Clark and Dudrick note, a neo-Kantian who accepts one of Kant's dualisms: the distinction between reason and nature. This is precisely the kind of dualism that a whole series of post-Kantian philosophers, from Schelling to Hegel and beyond, strove to overcome. So Nietzsche would have been aware of two traditions in philosophy: one, reaching its apogee in Kant, that distinguishes normative and causal realms; another, exemplified by Hegel, that dissolves this dualism. It seems to me decidedly odd to interpret Nietzsche as falling on Kant's side in this debate. Put a bit differently, it's surprising to think that Nietzsche would endorse a strong fact/value distinction—isn't he constantly emphasizing the ways in which allegedly non-evaluative, affectively neutral descriptions actually presuppose and contain evaluations? Where does one find Nietzsche countenancing such a distinction? To the extent that he writes about such a distinction, he seems to me extremely critical of it. Given Nietzsche's claim that even *sense-perception* is evaluative, it seems perfectly clear where he stands: not with Kant but Hegel.²²

Second, Clark and Dudrick make much of Nietzsche's rejection of "mechanism"; they take his rejection of mechanism to involve an endorsement of the causal/normative distinction. However, I think it's clear that Nietzsche thinks the mechanists go wrong not because they focus solely on the causal and ignore a separate normative realm, but because they don't see that our account of the causal is already, and pervasively, normative. The mechanists' conception of causes, substances, subjects, and so on already contains—or so Nietzsche wants to claim—evaluations.

Put a bit differently, Clark and Dudrick's distinction between the causal and the normative seems to presuppose a view of the following form:

- (A) We can offer a non-evaluative or non-normative description of the world, but if we do we will miss something crucial. The mechanists and other clumsy naturalists go wrong because their descriptions of the world leave out values and reasons.

²² "All experiences are moral experiences, even in the realm of sense-perception [*Es giebt gar keine anderen als moralische Erlebnisse, selbst nicht im Bereiche der Sinneswahrnehmung*]" (GS 114). See also D 119.

Whereas I suggest that Nietzsche's point is rather:

- (B) We cannot offer a non-evaluative or non-normative description of the world, because our descriptions ineluctably presuppose or constitute evaluations. The mechanists and other clumsy naturalists go wrong because they do not realize that their descriptions of the world, which purport to be purely factual, harbor or constitute evaluations.

Whatever we think of the tenability of the causal/normative distinction in general, it seems clear that Nietzsche rejects (A) and endorses (B). Put a bit differently, Nietzsche denies that there is any strong fact/value distinction—as Chapters Four and Five demonstrated, he constantly emphasizes the ways in which allegedly non-evaluative, affectively neutral descriptions actually presuppose and contain evaluations.

I conclude that there is no reason to saddle Nietzsche with the Kantian distinction between a realm of causes and a realm of reasons. Let's now consider the claim that drives could stand in normative relationships to one another. Clark and Dudrick want to claim both that drives recognize authority relations among themselves, and that we must distinguish recognition of authority from “mere physiological strength of the drives” and “brute causal strength” (2012, 198). Thus, when drive A commands drive B, this involves more than A simply overpowering B. Instead, drives issue commands when they exert “*political authority*” (2012, 198). Drive A commands drive B when drive B recognizes and responds appropriately to drive A's authority. Chapter Four discussed some significant problems with Clark and Dudrick's interpretations of drives. There, I argued that Clark and Dudrick's homuncular reading of drives is both philosophically and empirically problematic, and textually indefensible. The idea that drives stand in normative relationships—that they perceive authority relations, political standing, and so forth—relies on the homuncular interpretation.

So there are two problems with Clark and Dudrick's interpretation: Nietzsche would not accept Clark and Dudrick's distinction between a realm of causes and a realm of reasons; and Nietzsche would not claim that drives, considered in isolation from the whole agent, have the capacity to recognize normative relations. But suppose we set both of these problems aside, and assume both that Nietzsche would accept a normative/causal distinction and that drives can stand in political relationships to one another. Notice that even if we granted both of these points, Clark and Dudrick would encounter the problems mentioned above. For suppose that some drive—such as the sex drive—is recognized by other drives as standing in the highest position of political authority. Suppose further that the agent is an ascetic who views sex as disvaluable. When the agent acts in order to express the sex drive, he will presumably view himself as overcome by a part of himself. Just as the causally dominant drive needn't be accepted by the agent as expressive of himself, so too the normatively dominant drive needn't be accepted.

In response, Clark and Dudrick might claim that this is impossible: the normative order among the drives determines the agent's evaluative commitments. But this seems implausible: agents can fail to recognize the normative order among their drives, as

when the *Genealogy's* scientist fails to recognize that his ascetic drive enjoys a dominant position. Discrepancies between the normative order of drives and the agent's conscious thoughts about value must be possible, and insofar as they are, the scenario mentioned above can arise.

In sum, Clark and Dudrick's reading of the unified agent as the one who enjoys a particular normative ordering among the drives is problematic on at least three grounds: Nietzsche does not endorse a normative/causal distinction; the idea that drives recognize normative relationships is implausible; and, even if we set that aside and grant that drives can stand in these relationships, the account cannot make sense of discrepancies between what the agent values and which drive is "normatively" dominant.

7.3 Schiller on Unity as Harmony

On Plato's account, the alcoholic individual would be regarded as dominated by Appetite, rather than by Reason. Thus, Plato would have good reason to judge the alcoholic defective as an agent. This seems to be the right result. The suggested reading of Nietzsche, above, was supposed to be neutral on which drive dominated; unity was supposed to consist of domination by *any* drive. So the suggested reading cannot judge the alcoholic agent to be defective.

What Plato's model is capable of capturing, and the predominance model is not, is the fact that when the agent's reflective judgments conflict with the agent's predominant motive, we regard the agent as overpowered by a part of himself. Nietzsche himself would agree with Plato on this much. For, as we saw in Section 7.1, Nietzsche associates genuine agency with the ability to control one's behavior via choice. We are still attempting to explain exactly how Nietzsche understands these notions, but the unity as predominance model simply ignores them. So we need a different model.

At this point, it will be helpful to introduce another model of unity, which would have loomed large in Nietzsche's mind. Friedrich Schiller proposed a model of unified agency, which he conceived as an alternative to the Kantian model of agency.^{23,24} Nietzsche was familiar with Schiller's work,²⁵ and, like Schiller, took Kant's model of agency as a target and a point of departure. So it stands to reason that Nietzsche drew from Schiller's model.

²³ Schiller develops these ideas in several works, including his essay "On Grace and Dignity" and the *Letters on the Aesthetic Education of Man*. For the sake of brevity, I will confine my discussion to "On Grace and Dignity." References to this essay are in the following format: NA followed by page number refers to the pagination in Schiller's *Nationalausgabe*; the second page number refers to the English translation of the essay in Curran and Fricker (2005).

²⁴ Kant responds to Schiller in a footnote to *Religion within the Limits of Mere Reason* (6:23–24n). For a helpful discussion of what is at issue in the debate between Schiller and Kant, see Baxley (2003).

²⁵ According to Brobjer (2008, 254), Nietzsche read Schiller's *Asthetische Erziehung* in 1862, his ten volume *Sämtliche Werke* in 1871, and his *Über Naïve und sentimentalische Dichtung* in 1871 and 1873. The KGW contains at least 221 references to Schiller.

Schiller begins by accepting a version of the Kantian distinction between reason and sensibility. The individual, Schiller tells us, has two aspects: a rational nature, manifest in judgment and self-conscious thought, and a sensible nature, manifest in sense perception and affects.²⁶ These two aspects of human nature can be related in three different ways.²⁷

First, the individual might be dominated by his sensible nature, merely acting on whatever inclination happens to arise. “Prey to desire,” he “lets natural impulse rule him unrestrainedly” (NA 280/147). Schiller terms such an individual *ochlocratic* (i.e., ruled by a mob). Schiller denounces this type of individual, claiming that he is analogous to a failed state in which citizens do not acknowledge the legitimacy of their sovereign (NA 282/148).

Second, the individual might be dominated by his rational nature. Schiller takes Kant to endorse this state of the soul (NA 282–5/148–50). Schiller terms such an individual *monarchic*; his rational nature rules his sensible nature with “strict surveillance” (NA 281–2/148). Schiller claims that the monarchic agent is better off than the ochlocratic agent, for his actions will be in accordance with the balance of reasons, and will have moral worth.

Although the monarchic agent is superior to the ochlocratic agent, Schiller finds something problematic about *both* of these agents: namely, the fact that one *part* of the individual dominates the other part. “This much is clear: that neither the will . . . nor the affect . . . ought to use *force*” (NA 279/146). He endorses a third state: harmony between the rational and sensible parts of the soul. A harmonious individual would have affects that incline her to pursue the very same ends that rational thought inclines her to pursue. Like the monarchic agent, her actions would be in accordance with the balance of reasons. But unlike the monarchic agent, there would be no struggle, no antagonism, in the soul of this agent. Her whole being would incline her in one direction:

It is only when he gathers, so to speak, *his entire humanity* together, and his ethical way of thinking becomes the result of the united effect of both principles [e.g., Reason and Sensibility], when it *has become his nature*, it is then only that it is secure. (NA 284/150)²⁸

The human being has been set the task of promoting a sincere accord between his two natures, of always being a harmonious whole, and of acting with his whole harmonious humanity. (NA 289/154)

Schiller calls the harmonious individual the beautiful soul. Her actions have not only dignity, but *grace* (*Anmut*).

²⁶ Schiller claims that the human being’s “purely intellectual nature is accompanied by a sensuous one” (NA 284/149). He discusses this point at NA 257/128–9, NA 262/132, NA 266/135–6, NA 272–3/140–1, and NA 276–8/144–5.

²⁷ “One can think of three ways altogether in which a human can relate to himself, that is, in which the sensuous part can relate to the rational” (NA 280/147).

²⁸ Here I have departed from the translation in Curran and Fricker, which seems to me to obscure Schiller’s point. In the next passage, I have also made some minor modifications to the translation.

Thus, Schiller holds that an agent is unified when the two aspects of the soul—rational nature and affective nature—are harmonious, directing the agent toward the same ends. Disunity arises when there is a conflict between the rational and the affective, which takes the form of reason being out of accord with the affects. In short: unity obtains when the agent's reflective judgments and affects incline her in the same direction. Thus,

- Schiller's Parts: The self is divided into two parts: the rational and the sensible
- Schiller's Relation: The self is unified when the rational and sensible parts incline the agent toward the same ends.

There are two interesting parallels between Schiller's account and Nietzsche's account. Like Nietzsche, Schiller denies that unity can be achieved via one part's dominance of the other part(s). Moreover, Nietzsche describes his paradigm of unity (Goethe) in terms that are reminiscent of Schiller:

What [Goethe] wanted was *totality*; he fought against the separation of reason, sensation, feeling, and will [*das Auseinander von Vernunft, Sinnlichkeit, Gefühl, Wille*] (preached with the most abhorrent scholasticism by Kant, Goethe's antipode); he disciplined himself to wholeness. (TI IX.49)

Nietzsche tells us that Goethe fought against the "separation" [*das Auseinander*] of reason, sensation, feeling, and will. The reference to Kant suggests that Nietzsche is making Schiller's point: Kant allegedly thought that reason should dominate passion. Nietzsche, with Schiller, conceives a harmony between the various aspects of the soul. Goethe exemplifies that state.

So Nietzsche seems to accept a version of Schiller's idea, that unity is attained when the parts relate to each other in a harmonious way, rather than when one part dominates the other parts. In addition, I will suggest that Nietzsche accepts and develops Schiller's basic point: the agent is unified when there is a harmonious relationship between the agent's reflective thought and the agent's affects. Notice that, in the passage quoted above, Nietzsche is *not* referring to a unity among drives alone. Rather, he speaks of attaining unity among *reason, sensation, feeling, and will*. Now, I think it would be a mistake to put too much weight on the particular terms that Nietzsche is using. For example, we should not interpret Nietzsche as claiming that the self has exactly four discrete parts that must be unified, for Nietzsche elsewhere rejects claims of this form. However, it is interesting that Nietzsche speaks of a unity not among drives alone, but among drives and other parts of the self. This helps us to make sense of Nietzsche's claim that unity is compatible with conflict among drives (see section 7.2.2). If Nietzschean unity were unity amongst drives, then Nietzsche's praise of conflict amongst drives would be odd. However, if Nietzschean unity were unity between, say, drives and reason, then conflict amongst drives need not be relevant. Take an analogy: Plato does not worry about the fact that the part labeled Appetite contains, within itself, many conflicting particular appetites. *That* kind of

conflict does not endanger unity, because Platonic unity is unity between Appetite, Reason, and Spirit. Just so, I suggest, Nietzsche is not inconsistent when he praises individuals with conflicting drives, for Nietzschean unity is *not* unity between particular drives, but unity between drives and other parts of the individual.²⁹

What are these parts? I suggest that they are roughly the ones that Schiller focuses upon: Reason and Sensibility. Of course, Nietzsche would not characterize the parts in these terms, nor would he agree with Schiller on the exact nature of these parts. However, Nietzsche draws attention to a closely related state. This is the state in which the agent experiences a particular kind of discrepancy or discord between his reflective judgments (Reason), on the one hand, and his drives and affects (Sensibility), on the other hand. To explain this point, I am going to spend some time examining one of Nietzsche's paradigms of disunity: the priests of the *Genealogy*.

7.4 An Analysis of Nietzschean Unity

In this section, I argue that we can grasp Nietzsche's concept of unity by examining a paradigm of disunity: the *Genealogy's* ascetic priest.³⁰ I argue that Nietzschean unity consists in a certain relation between an organism's reflective and unreflective aspects at the time of action. Although many commentators have assumed that Nietzschean unity is a unity between drives, I argue that there are persuasive textual grounds for interpreting Nietzschean unity as a unity between drives and self-conscious thought.³¹

7.4.1 The Character Type "Priest"

The First Essay of the *Genealogy* introduces three "character types": the slave, the noble, and the priest. Although my goal is simply to characterize the psychic state of the priest, describing this psychic state requires a brief reconstruction of the main argument in

²⁹ I do not mean to suggest that proponents of the predominance model have no way of explaining Nietzsche's praise of conflict amongst drives. On the contrary, they do have ways of doing so. For example, one could argue that Nietzsche wants *both* the predominance of one drive *and* constant struggle amongst the dominated drives for predominance. There is nothing incoherent about this proposed solution. However, my suggested interpretation has the advantage of rendering these sorts of proposed solutions completely unnecessary. For, on my interpretation, there is no *problem* to solve: Nietzsche's praise of conflict amongst drives only looks puzzling and problematic if we assume that the unity he desires is a unity amongst the drives. So my proposed solution has the advantage of *dissolving* the apparent problem, thereby rendering a quest for solutions unnecessary. Notice that if my reading is correct, it has the added benefit of explaining why Nietzsche himself never seems to detect any potential tension between his endorsement of agential unity and his valorization of conflict amongst the drives.

³⁰ Nietzsche frequently emphasizes the priest's disunity. The ascetic priest is characterized as "a self-contradictory type" (GM III.11), manifesting "inner corruption" (GM III.14) and "ill-constitutedness" (GM III.13). Nietzsche further criticizes the priest for being "discordant," a "self-contradiction" (GM III.13). He is the "shepherd" of the "ill-constituted" (GM III.13); he seeks to make the "well-constituted" suffer (GM III.14), he "walk[s] among us as embodied reproaches, as warnings to us—as if health, well-constitutedness, strength... were in themselves necessarily vicious things" (GM III.14).

³¹ This fact helps to explain why Nietzsche treats disunity as a condition peculiar to self-conscious animals. A non-self-conscious animal, on Nietzsche's view, cannot be disunified.

the First Essay.^{32,33} I should note that the First Essay's argument is highly complex, and there are a number of controversies surrounding the structure of the argument. For our purposes, nothing important hangs on the precise way in which the pieces of the argument fit together, so I will pass over these controversies, simply mentioning them in the footnotes.

Nietzsche begins by juxtaposing two types of individuals: the heroic, strong, healthy nobles, and the weak, craven, downtrodden slaves. Initially, both types of individuals accept a system of values that posits health, strength, and beauty as the chief goods, and sickness, weakness, and ugliness as the chief forms of badness. By the light of these values, the slaves are bad, and the nobles good. Thus, the slaves are committed to regarding themselves as bad, unfortunate wretches. Moreover, they do not see this as a temporary or surmountable state of affairs; rather, they view themselves as irredeemably bad (GM I.6).

Nietzsche next introduces a third type of individual, the priest (GM I.6–7, I.16). Like the slaves, the priests are weak and unhealthy. Unlike the slaves, the priests are not content to resign themselves to this state of affairs. On the contrary, the priests have a “lust to rule,” an ardent desire to occupy positions of power and influence (GM I.6). Thus, they are engaged in a struggle for dominance with the nobles (GM I.6–10). However, the priests have none of the traits that are regarded as valuable: they are not physically strong, healthy, and so forth (GM I.6–7).

Frustrated by their inability to attain dominance, the priests come to bear an intensely negative affect toward the nobles, an affect that Nietzsche calls *ressentiment*. *Ressentiment* has several distinctive features. First, and most obviously, *ressentiment* involves a negative affect of hatred and vengefulness (GM I.10, III.14). This hatred is directed at the nobles (GM I.10). It is crucial that the priest hates the noble for a particular reason: he wants to possess the characteristic traits of the noble—he wants to be beautiful, strong, wealthy, and healthy—but finds himself completely incapable of doing so. In other words, the priest finds himself unable to realize the form of life that he regards as most valuable, and is able to live only in ways that he regards as disvaluable. Confronted with those who enjoy the form of life that he values and ardently desires, the priest comes to hate the nobles. But notice that this intense hatred is directed at those who, by the priest's own lights, are living the good life; moreover, the priest hates them precisely *because* they are leading the good life.

So the priest has an odd response to the noble: he values their way of life, and for that reason it would make sense for him to regard the nobles with admiration or respect; but, instead, the priest hates the nobles. This creates a discord between the priest's reflective evaluations and his emotions.

³² The slave also provides an excellent example of disunity. However, I have chosen to focus on the priest, for the psychic tensions that I wish to highlight are somewhat clearer and more pronounced in the priest.

³³ My account of the First Essay draws on Reginster (1997) and Jay Wallace (2007).

Typically, emotions and evaluative judgments are reciprocally related. For emotions involve a claim to rationality: they have standards of intelligibility.³⁴ Consider an example. Hating a person commits one to viewing the person as in some respect bad. There are borderline cases: you can, I suppose, hate someone for no reason, though this verges on unintelligibility. But it does not make sense to hate someone whom you regard as having wholly positive attributes, for hating a person involves judging the person to be in some respect disvaluable.

The priests violate this standard. The nobles embody all the traits that the priests reflectively evaluate as good, yet the priests have an intensely negative affect directed at the nobles.

This kind of discord between affect and evaluative judgment can be eliminated in two ways. The agent can attempt to change his affects, bringing his affects into accordance with his evaluative judgments. Or the agent can attempt to change his evaluative judgments, bringing his evaluative judgments into accordance with his affects. The priests take the second route. As Jay Wallace puts it:

The powerless find themselves in a conceptual situation in which the negative affect that dominates their lives is directed at individuals whom they themselves seem compelled to regard as exemplars of value and worthy of admiration. This is a highly unstable combination of attitudes... In my view, the slave revolt should be understood as a response on the part of the slavish to this psychic tension. The weak are subject to attitudes that color their experience of the social world, in ways that cannot be reconciled with the dominant ethical ideology that they themselves have so far accepted. So they come to embrace a new and more congenial scheme of values. (Wallace 2007, 220)³⁵

This new set of values labels health, strength, and beauty as evil, and sickness, weakness, and ugliness as good (GM I.7). In effect, it inverts the older system of values. By the lights of this new set of values, the noble is evil. So the priest's *ressentiment* makes sense: if the noble is evil, then the priest is justified in bearing hatred or *ressentiment* toward the noble.³⁶

³⁴ See, for example, D 119, 279; GS 14, 39; Z I.15; BGE 187, 220, 230, 260, 268, 284; GM II.16, III.12; TI IX.20. For a very helpful discussion of Nietzsche's views on this point, see Peter Poellner (2007).

³⁵ Wallace's language is misleading: Nietzsche certainly would not be troubled by the fact that an agent bears "a highly unstable combination of attitudes." After all, Nietzsche explicitly praises those who bear contradictory, unstable affective states. For example, he writes that the individual with "contradictory drives" has "a great method of acquiring knowledge: he feels many pros and cons... The wisest man would be the one *richest in contradictions*" (KSA 11:26[119]; cf. HH 618, GM III.12, BGE 284). Thus, Wallace's concerns about unstable combinations of attitudes seem out of place. However, I think the idea that Wallace is attempting to express with his claims about "unstable attitudes" is exactly right: as I will explain below, the problem with the priests is that they exhibit a discrepancy between affects and evaluative judgments.

³⁶ The values preached by the priests come to be accepted by the slaves. It is easy to see why: so long as the slaves embrace the older system of values, they are committed to regarding themselves as bad. Thus, they will experience shame and frustration at their current condition. If they accept the new system of values, they will regard themselves as good. Under the influence of a desire to think well of themselves, the slaves come to accept the priestly values. This is what Nietzsche terms the "slave revolt" in values: the new system of values takes root among the slaves.

Let us summarize the argument so far. The priests come to bear a negative affect of *ressentiment* toward the noble. This creates a state of discord between affect and evaluative judgment; the priests have an intensely negative affect directed at those who epitomize their own values. The priests respond to this state of tension by altering their evaluative judgments.³⁷

There is one more important complication. Recall that the priest came to resent the noble because the priest was engaged in a struggle for dominance with the noble. That is, the priest wanted to occupy positions of political power, wanted to dominate, and so forth. After the revaluation, the priest judges traits such as the desire for dominance to be evil. However, Nietzsche makes it clear that the priest continues to desire to be dominant. Indeed, his very process of revaluing values was undertaken because he ardently desired dominance. As Reginster puts it, “the priests who so vehemently condemn the thirst for ‘spoil and victory’ of the noble ‘blond beast’ (GM I.11) are in fact pursuing the very same ‘goals . . . victory, spoil, and seduction’ (GM I.8)” (Reginster 1997, 291).

So although the revaluation eliminates one form of discord between affect and evaluative judgment, it gives rise to a *second* form of discord between affect and evaluative judgment: the priest ardently desires that which he judges evil. Confronted with the discrepancy between affect and evaluative judgment, the priest resorts to self-deception:

The man of *ressentiment* is neither upright nor naïve nor honest and straightforward with himself. His soul *squints*. (GM I.10)

[The man of *ressentiment*] has, thanks to the counterfeit and self-deception of impotence, clad itself in the ostentatious garb of the virtue of quiet, calm resignation, just as if the weakness of the weak . . . were a voluntary achievement, willed, chosen, a *deed*, a *meritorious* act. (GM I.13; cf. I.14)

Through self-deception, the priest hides from himself the fact that he desires that which he condemns as evil.

7.4.2 Important Features of the Character Type

Notice that the priests exhibit *precisely* the structure that Schiller identifies with disharmony: their reflective judgments are out of accordance with their affects; the two

³⁷ This interpretation is somewhat controversial. A different, and perhaps more familiar, interpretation claims that the priests engage in revaluation not in order to eliminate their own psychic tension, but in order to succeed in their struggle for dominance with the nobles. On this interpretation, the priests believe that if they can convince the slave class that the nobles are evil, then the priests will be able to assume a dominant position. So the priests engage in revaluation in order to fulfill their desire for dominance. I think this interpretation faces a number of difficulties, which are discussed by Wallace (2007). However, nothing that I am about to say about the psychic state of the priest hinges on this point. Readers who prefer the second interpretation of the priestly revaluation can substitute this interpretation for the one offered in the text, above. For my purposes, the only important point is that the priests have a form of discord between affect and evaluative judgment. It does not matter whether this discord motivates the revaluation, or plays a less substantial role.

parts direct the agent toward opposing ends. As Schiller would put it, the rational and the sensible parts of the self are in conflict.

This discord between affect and evaluative judgment shows up at two crucial junctures of the *Genealogy*. First, when the priest begins to bear *ressentiment* toward the noble, the discord between affect and evaluative judgment motivates the priest's revaluation of values. Second, after the revaluation, the priest experiences discord between his ardent desire for dominance and power, on the one hand, and his reflective condemnation of those traits, on the other hand. This discord motivates the priest's self-deception.

Given the centrality of this notion, we might be tempted to identify it with Nietzschean disunity. That is, we might claim that an agent is disunified iff the agent's reflective judgments conflict with the agent's drives and affects, in the sense that they direct the agent toward conflicting ends. I think this is almost, but not quite, correct. There is an additional facet to Nietzsche's account of disunity: discord between reflective judgment and affect is not sufficient to engender disunity. It is only *certain forms* of conflict between reflective judgment and affect that lead to disunity.

There are two pieces of textual evidence indicating that certain forms of discord between affect and reflective judgment are compatible with agential unity. First, Nietzsche often enjoins us to put ourselves in states of discord between affect and evaluative judgment. He claims that conflicts between affects and evaluative judgments are an excellent opportunity for acquiring a deeper understanding of our values. He writes that the individual with "contradictory drives" has "a great method of acquiring knowledge: he feels many pros and cons . . . The wisest man would be the one richest in contradictions" (KSA 11:26[119]; cf. HH 618, GM III.12, BGE 284). Indeed, Nietzsche describes his genealogies as operating in precisely this way: they are meant to generate an emotional reaction to the value under investigation, thereby creating a discord between the affective response to valued object and the reflective evaluation of the object.³⁸ For example, a genealogy of pity might proceed by fostering a negative emotional response to pity, in order to generate a discord with our positive evaluation of pity. This state of tension leads us to reassess the value (of course, this reassessment need not culminate in the rejection of the value).

Now, the fact that Nietzsche enjoins us to enter these discordant states does not by itself indicate that conflict between affect and evaluative judgment is compatible with agential unity. After all, we could identify disunity with conflict between affect and evaluative judgment, and interpret Nietzsche as claiming that disunity is instrumentally valuable, as a means to some valued end. But recall that unity is meant to be an account of agential activity. It is hard to believe that the individual who engages in a

³⁸ Nietzsche writes, "the inquiry into the origin of our evaluations and tables of the good is in absolutely no way identical with a critique of them, as is so often believed: even though the insight into some *pudenda origo* certainly brings with it a feeling of diminution in the value of the thing that originated that way and prepares a critical mood and attitude toward it" (KSA 12:2[189]/WLN 95).

Nietzschean genealogy, and discovers formerly hidden aspects of her psychic states and values, thereby becomes *less* active in the production of her actions.³⁹

This brings us to the second piece of textual evidence. Some of the individuals whom Nietzsche praises—especially, Goethe and Nietzsche himself—are praised for harboring and tolerating inconsistent values:

Human beings have in their bodies the heritage of multiple origins, that is, opposite, and often not merely opposite, drives and value standards that fight each other and rarely permit each other any rest. Such human beings of late cultures and refracted lights will on the average be weaker human beings: their most profound desire is that the war they *are* should come to an end... But when the opposition and war in such a nature have the effect of one more charm and incentive of life—and if, moreover, in addition to his powerful and irreconcilable drives, a real mastery and subtlety in waging war against oneself, in other words, self-control, self-outwitting, has been inherited or cultivated too—then those magical, incomprehensible, and unfathomable ones arise, those enigmatic men predestined for victory and seduction, whose most beautiful expression is found in Alcibiades and Caesar... and among artists perhaps Leonardo da Vinci. They appear in precisely the same ages when that weaker type with its desire for rest comes to the fore: both types belong together and owe their origin to the same cause. (BGE 200; cf. BGE 224)

Nietzsche claims that having conflicting drives and “value standards” often *leads to* disunity: it leads to “war,” “opposition,” and so on. However, certain individuals remain unified despite experiencing these forms of conflict: these individuals manage to achieve “self-control.” In other words, disunity is a frequent result of having discordant affects and evaluative judgments, but disunity is not identical to having discordant affects and evaluative judgments.

I think these two pieces of textual evidence indicate that Nietzsche treats discord between affect and evaluative judgment as a condition that frequently results in, but is not identical to, disunity. Setting the textual evidence aside, though, there is also a compelling philosophical reason for distinguishing disunity from discord between affect and evaluative judgment. Recall that the concept of unity is designed to play a particular philosophical role in Nietzsche’s account of agency: when an agent is unified, the agent plays an active role in producing her action; when she is disunified, her action is produced independently of her participation. Nietzsche quite rightly points out that having conflicting affects does not undermine an action’s attributability to the agent.⁴⁰ After all, it is possible for an agent to harbor massively conflicting sets of affects, while remaining in control of her actions. Just so, the mere fact that an agent has conflicting affects and evaluative judgments does not indicate that the agent plays no role in the production of her action. For it is possible, albeit difficult, to harbor conflicting affects and evaluative judgments while remaining in control of one’s actions.

³⁹ And indeed, Nietzsche seems to associate increasing self-knowledge with increasing activity. For an illuminating discussion of this point, see Richardson (2004).

⁴⁰ See especially KSA 11:27[59] and BGE 212, quoted above on page 176 in Section 7.2.2.

Consider an example of an individual who experiences a conflict between his affects and his values. Earlier, we considered an alcoholic who ardently craves a drink, but judges that drinking would be disvaluable. Suppose this agent experiences frequent and powerful desires for alcohol. Yet he values sobriety and disvalues alcohol. His affects and values are discordant, for the desire for alcohol conflicts with the disvaluation of alcohol. This agent is certainly in an unfortunate state; it will be hard for him to control his actions, and he may end up succumbing to the desire for alcohol. But suppose he does manage to avoid drinking. Rather than drinking the alcohol, he pours it down the sink. That action—pouring the alcohol down the sink—seems attributable to the agent's activity. The fact that the action is rendered difficult by opposing affects, which tempt the agent to refrain from pouring out the alcohol, does not seem to show that the agent is passive in the production of his action. If anything, the opposite is the case. Put simply, the discordant affects make it hard for the agent to control the action, but do not render the action not his own.

7.4.3 Nietzschean Unity

In the prior section (Section 7.4.2), I pointed out that the priests exhibit a conflict between their reflective judgment and their affects. I argued that this feature does not, by itself, engender disunity: certain agents display the same type of conflict, but remain unified. So what differentiates the priests from these unified individuals?

Notice that the conflict between reflective judgment and affect takes a special form in the priests: it is *hidden*. I will argue that this fact is important.

Of course, there is nothing unusual about being ignorant of various aspects of one's actions. Nietzsche disparages the idea that agents are generally cognizant of their actions, calling it "the universal madness." For "the opposite is precisely the naked reality demonstrated daily and hourly from time immemorial! . . . Actions are *never* what they appear to be . . . all actions are essentially unknown" (D 116; cf. D 119).

However, the priests have a distinctive form of self-ignorance. I think we can capture their condition as follows:

(Disunity) The agent currently approves of his A-ing. However, if he knew more about the drives and affects that figure in A's etiology, he would *not* approve of his A-ing.

This notion of disunity is meant to characterize a peculiar psychological state, in which an agent approves of his action, but would reverse the attitude, were he to know more about the action's causal history. (The particular elements in the causal history that occupy Nietzsche's attention are the agent's drives and affects, though in principle there is no reason for restricting it to these types of motives.)

Notice that disunity, so defined, constitutes a kind of dissatisfaction with one's action. If an agent is disunified, he would cease to approve of his action, were he to know more about its etiology. So we might state the definition of disunity as follows: an agent is disunified in performing an action A if, were he to know more about the drives and affects that are causing him to A, he would *not affirm* A-ing.

With this in mind, let us offer a characterization of the contrary of disunity, namely *unity*:

(Unity) The agent *A*s, and affirms his *A*-ing. Further knowledge of the drives and affects that figure in *A*'s etiology would not undermine this affirmation of *A*-ing.

The account should be understood as applied to agents, holding all else constant except giving the agent further information about the drives and affects figuring in the etiology of the action under consideration. In particular, we do not want to consider cases in which the agent changes his values. For example, I would now disapprove of many of the actions that I performed, approvingly, as a child; but this does not show that I was disunified in performing the actions, for at the time of action I may have wholeheartedly approved of them.⁴¹

Additionally, notice that the only factor that we are changing here is how much information the agent has about the *etiology* of the action. Agents sometimes disapprove of a past action, not because they learn more about the act's etiology, but because they learn more about the act's *consequences*. I did not realize that my innocent, offhand remark would hurt Sarah's feelings. Now, seeing her upset, I regret the remark, and wish that I had not made it. But this does not show that I was disunified in acting.

Notice, finally, that a disunified agent would not necessarily want to *act* differently. She might be dissatisfied with her actions, not because she disapproves of what she has done, but because she disapproves of her *motives* for doing what she has done. For example, suppose that Sally volunteers in a soup kitchen. She believes she is volunteering out of a desire to aid the impoverished beneficiaries. Yet a psychologically adept observer, well acquainted with Sally's character, would describe things differently: Sally takes satisfaction in feeling superior to the impoverished recipients, and her volunteering is in part motivated by this desire. Suppose Sally comes to realize that one of the desires motivating her action is the desire to feel superior. She finds this desire reprehensible, and she is no longer able to view her action of volunteering with approval. Thus, she is disunified. However, it would be inaccurate to say that she wants not to volunteer. Rather, she still wants to volunteer, but she wants to volunteer out of beneficent motives, rather than self-serving ones. So she is disunified, not because she wants to act differently, but because she wants to act out of different motives.

With these points in mind, we can see that disunity constitutes a form of psychic conflict. An agent acts, and approves of his action. However, this approval is contingent upon ignorance of the drives and affects that are actually leading him to act. So there is

⁴¹ Nietzschean unity bears a resemblance to Harry Frankfurt's notion of wholeheartedness. Roughly, Frankfurtian wholeheartedness obtains when the agent bears a higher-order attitude of acceptance or approval toward his lower-order desires (see Frankfurt 2004). Nietzschean unity is more demanding: the approval in question must be stable in the face of further information about the action's etiology. There is a striking similarity between this view and Christman's (1991) account of freedom; thanks to Ken Baynes for alerting me to this similarity.

a conflict between the agent's attitude toward the action as he takes it to be, and the agent's attitude toward the action as it is. Moreover, disunity implies that one has affects and drives that are moving one in ways that one would disavow. Thus, there is an interesting form of conflict between the agent's reflective and unreflective aspects at the time of action.

The notion of disunity can be used to characterize the psychic state of the priests. The priests are ignorant of aspects of their actions; they do not realize that they lust after the very states of affairs that they reflectively condemn. To adopt Nietzsche's pithy formulation, "the *motives* of this morality stand opposed to its *principle*" (GS 21). Section 7.4.2 argued that this discord between affect and reflective judgment is compatible with the agent's playing an active role in the production of his action. The struggling alcoholic, who successfully overcomes his cravings for alcohol, plays an active role in the production of his action, despite his experiences of discord. However, notice that the priest's discord is *hidden*. Rather than struggling with this discord in the open, or attempting to resolve the discord, the priest hides the discord from himself: the priest is "neither upright nor naïve nor honest and straightforward with himself" (GM I.10). This combination of discord and self-ignorance has an interesting result: it renders the priest disunified.

Consider an action that a priest performs. Say, he inveighs against the noble for being powerful. The priest's moral system gives power a negative evaluation. Thus, the priest will approve of his philippic. However, the priest has a "lust to rule"; he, too, seeks power. If he were to discover that this attitude were present in the etiology of his action, the priest would not be able to approve of his action. Thus, he is disunified.

In sum, then, I submit that Nietzsche's analysis of unity is as follows:

- Nietzschean Parts: the self can be divided into drives and affects, on the one hand, and reflective thought, on the other hand.
- Nietzschean Relation: the self is unified when the reflective and unreflective parts of the agent are harmonious, in the following sense: the agent A's, affirms his A-ing, and further knowledge of the drives and affects that figure in A's etiology would not undermine this affirmation of A-ing.

When the agent manifests this form of unity, she produces a genuine action.

7.5 Is Nietzschean Unity an Adequate Account of Genuine Agency?

We now have an analysis of what Nietzsche's conception of unity is. But we still need to ask whether the concept fulfills its philosophical role, namely, distinguishing genuine actions from their lesser relatives. To answer this question, it will help to examine a case in which an agent comes to recognize his own disunity. Consider a famous quotation from Augustine, in which Augustine describes his effort to extirpate his attraction to pride:

Even when I reproach myself for it, the love of praise tempts me. There is temptation in the very process of self-reproach, for often, by priding himself on his contempt for vainglory, a man is guilty of even emptier pride. (Augustine, *Confessions*, Book X, 38)

Augustine is endeavoring reflectively to revalue pride. Formerly, he valued pride; now, he regards it as evil. Yet the revaluation is no easy task: despite his efforts, he finds himself entangled in lingering affective traces of the old valuations. His desire to reproach himself for instances of pride is motivated, in part, by his persistent attraction to pride. He prides himself on reproaching himself for being prideful. So the renounced value lives on, in the sense that Augustine continues to have positive affective responses to pride.⁴²

Here we have an individual who harbors a kind of psychic tension: the original evaluation of a state of affairs lingers on in a motivational tendency that opposes the motivational tendency springing from the new and different evaluation of the state of affairs. We can put the point this way: Augustine reflectively revalues pride, but his affects lag behind; he continues to bear the affects that were associated with the valuation of pride (e.g., attraction to pride, delight in instances of pride), and these affects mingle with and influence the affects that are associated with disvaluation of pride (e.g., shame at displays of pride, aversion to pride). Thus, Augustine manages to take delight in his shame at displays of pride, thereby manifesting a mixture of the old and new affects.

Like the *Genealogy's* priest, Augustine experiences discord between his affects and his reflective judgments. Unlike the priest, Augustine strives to recognize and eliminate the manifestations of this discord. He often fails: above, he describes his discovery that his reflective assessments of his own motives are being influenced by the very motives that he is attempting to disavow. That is, he discovers that his reflective disavowal of pride is motivated by pride. So, in reflectively disavowing pride, he is disunified: once he learns more about the etiology of these disavowals, he comes to disapprove of them. For he can no longer see these disavowals of pride as expressive of his disvaluation of pride; on the contrary, he sees that these disavowals of pride are being motivated by his persistent attraction to pride.

With this discovery, Augustine is revealed to have been less than fully active in the production of his past actions. For his choices were being influenced in a fashion that he did not recognize, and which he would have found objectionable. This shows that Augustine's action was not wholly the product of his own activity.

Like the *Genealogy's* priest, Augustine is strongly attracted to that which he reflectively condemns. Moreover, Augustine does not initially notice this attraction, so he is initially disunified. And the problem that Augustine's disunity produces is that there is no clear answer to the question of where the agent stands. He has chosen to denounce pride, and this choice is consistent with his evaluative judgments. However, the denunciation is motivated by an attitude that Augustine cannot condone: pride itself. So

⁴² See Reginster (1997, 293) for a discussion of the relevance of this passage for Nietzsche.

Augustine's condemnation of pride no longer appears to be a manifestation of agential activity. His reflective thought seems to be buffeted about by inconsistent drives and affects, in a way that renders him a passive conduit for forces within.

As this example indicates, Nietzschean unity is at least a necessary condition for agential activity. Unity seems to offer a characterization of the conditions under which an agent can be said to be in control of her action.⁴³ The agent acts, approves of the act, and further knowledge of the action would not undermine this approval. To speak metaphorically, the agent's whole being is behind the action. I conclude that Nietzsche's distinction between unity and disunity adequately captures the distinction between genuine action and mere behavior.

Moreover, notice that Nietzsche's account of unity gives us a way of distinguishing genuine cases of agency from their lesser relatives without committing ourselves to concepts of the will or the self that Nietzsche would regard as problematic. In order to manifest unity, an individual need not be pictured as standing apart from her drives and exerting a controlling influence on them. Nor must we conceive of the agent as having and exercising a will that is wholly independent from her drives. Rather, the account of unity merely requires that agents have conscious thoughts and engage in episodes of deliberation and choice, on the one hand, and also have drives and affects, on the other. The conscious thoughts and the capacity for choice are pervasively and inescapably influenced by drives. Yet they are distinct from drives, and therein arises the potential for disunity.

To put this in terms of Chapter Six: this account of unity fits perfectly with the vector model of the will. Drives, affects, and conscious thoughts all exert their influences on forthcoming actions. The agent manifests unity when this tangle of forces exhibits a certain relation: unity, as described above.

7.6 Conclusion

Nietzsche's account of agential unity is designed to offer a way of distinguishing genuine manifestations of agency from acts that are produced merely by the agent's drives and affects. In other words, the account of unity distinguishes genuine actions from mere behaviors. The account of unity enables us to draw this distinction without relying on notions of the will or the self that Nietzsche would regard as problematic. Unity simply requires a certain kind of harmony between the agent's reflective and unreflective aspects at the time of action. In particular, an agent is unified—or, equivalently, the agent is active in the production of her action—when she approves of her action, and further knowledge of the action's etiology would not undermine this approval.

Moreover, Nietzschean unity is an analysis of what it is for the agent to determine her action through choice. Everyone will agree that *one* way of failing to

⁴³ Notice that an individual who is unified cannot act akratically. If an agent performs an action of which he disapproves, he is not unified.

determine one's action by choice is for one's action not to conform to one's choice: I decide not to drink at the party, but end up drinking after all. But Nietzsche draws our attention to another way that one can fail to determine one's action: one's *choice* can be determined by one's motives in a way that one would disavow were one to recognize it. Thus, Nietzsche points out that in addition to examining the connection between choice and action, we must examine the connection between *the agent* and *choice*.

8

Self, Culture, and Society

So far, we have seen that Nietzsche endorses a vector model of the will, according to which conscious thoughts, deliberations, and choices do not enjoy any privileged role in determining action. Instead, they are merely one component of the vector of forces leading to action. Nevertheless, in some cases conscious deliberation does crucially shift the balance of forces and alter the character of forthcoming actions. Many philosophers would want to claim that it's only in these kinds of cases—cases in which the conscious thoughts play a role in the production of behavior—that the agent counts as genuinely acting. But in Chapter Seven I argued that Nietzsche rejects this view. The agent's conscious deliberations and willings may be buffeted about by forces of which she is ignorant, in ways that would undermine her attitude toward her action were she aware of them. In light of this, Nietzsche introduces a notion of unified agency. When the agent exhibits this form of unity, she counts as active in the production of her action and there is a sense in which she can be said to choose her action. In other cases, where the agent exhibits disunity, she is better regarded as a passive conduit for forces within.

With this, we have completed our examination of Nietzsche's views on mental activity, motivation, value, and agency. We've seen how Nietzsche analyzes the conscious/unconscious distinction; how he introduces a distinctive psychological state called the drive; and how these analyses inform his account of motivation and agency. What remains to be done is to see how Nietzsche draws these points together in his analysis of the human person. What does this account of the psychological constituents of the person tell us about the Nietzschean self?

Let me clarify this question. At this point, we understand the reciprocal relations between unconscious and conscious states, between motives, values, and episodes of willing. In short, we understand what we might term the *internal* constituents of the self. But we haven't yet seen how the agent's social environment impacts these internal constituents. Human beings aren't isolated agents: we inhabit a realm of custom, convention, evaluative claims, norms, and expectations. Just as we've asked how the unconscious interacts with the conscious—how, at the most general level, the constituents of the human mind interact with one another—so too we can ask how the mind interacts with its social environment. And this is now our topic.

I begin, in Section 8.1, by pointing out that Nietzsche treats selfhood as an aspirational term: we are not selves merely in virtue of being human. Rather, Nietzsche

claims that selfhood is something that must be attained. But what kind of attainment might selfhood be?

Section 8.2 considers one answer to this question: a person counts as a self when her drives or actions cohere, directing her toward some overarching goal. This is by far the most common interpretation in the secondary literature, but I argue that it is mistaken: while Nietzsche does value some forms of coherence, he treats them as having only instrumental value. Coherence might contribute to genuine selfhood, but it is not sufficient for it.

In particular, I argue that Nietzsche uses the term “self,” as well as a cluster of related concepts (person, sovereign individual, great individual, higher individual, and so on), to refer to those who bear an appropriate relation to their culture. An analogy: just as really acting, really being an agent, requires a certain relation between unreflective and reflective components of the agent, so too being a self, being a *real* self, requires a certain relation between the person and her culture.

The idea that “self” and related terms pick out individuals who stand in appropriate relations to their culture is relatively clear in Nietzsche’s texts. Much more difficult, though, is determining exactly what the person/culture relationship is supposed to be. The remainder of the chapter investigates this topic.

Section 8.3 suggests that we can learn what kind of self–culture interaction Nietzsche is envisioning by examining some of his paradigm cases of great individuals. Nietzsche presents historical figures such as Goethe, Napoleon, and Beethoven as genuine selves. I suggest that part of what Nietzsche praises, here, is the great individual’s embodiment of a new ideal or promotion of a new set of “higher values.” The individual counts as great, by Nietzsche’s lights, when he not only embodies this new ideal, but also plays a transformative role, shifting groups or entire societies toward new hierarchies of value.

While greatness, so described, is sufficient for selfhood, Nietzsche does not present it as necessary; we can be genuine selves without being great individuals. So I ask, in Sections 8.4 and 8.5, whether one characteristic of the great individual—embodying a new ideal, or reassessing values, or providing new ways of life—might by itself be enough for genuine selfhood. I think that it is, but the analysis has to be conducted carefully: it’s easy to confuse Nietzsche’s account with superficially similar, yet problematic, accounts. Section 8.4 investigates one such account: the Romantic claim that the genuine self is the person who frees himself from the constraints of culture and returns to some essential, pre-social set of values or motivations. Though some commentators have been tempted to interpret Nietzsche along these lines, I argue that Nietzsche dismisses both the idea of pre-social drives and the assumption that we can characterize what a person is without making reference to determinate aspects of her social setting.

Section 8.5 suggests that a more promising approach would be to treat genuine selfhood as requiring critical assessment of one’s values. Yet this must be distinguished from a position that Nietzsche clearly rejects: the Kantian account of autonomous self-legislation. I introduce the problems with this Kantian picture in Section 8.5. The resolution occupies Chapter Nine.

8.1 Nietzsche on the Self

The first step in understanding Nietzsche's account of the self is making sense of two striking claims: that the self should be understood as a relation of drives, and that most of us lack selves. In the following sections, I investigate these claims in turn.

8.1.1 *The Self as a Structure of Drives*

Nietzsche repeatedly claims that the self is constituted by its drives. For example, he writes that the self is a "social structure of drives and affects" (BGE 12) and he tells us that the intellect "is actually nothing but a *certain behavior of the drives towards one another*" (GS 333). As I've indicated in the previous chapters, commentators often misinterpret this point, arguing that Nietzsche aspires to reduce the self to nothing more than relations of drives. This is a mistake. To say that the self is constituted by drives, or that the intellect's operations are to be understood as a relation of the drives toward one another, is not to say that the human organism is *nothing but* drives. Drives don't have eyes, they don't walk, they don't reflect, they don't talk. Drives are dispositions that token evaluative orientations, and Nietzsche's claims about drives constituting organisms should be understood as claims about how drives impact the organism's cognitive, perceptual, and motivational states. In other words, Nietzsche is emphasizing that drives are intertwined with agents' affects as well as their conscious thoughts and reflections, structuring large swathes of the mental economy. Thus, rather than interpreting Nietzsche as making reductive claims about the self comprising nothing but drives, we should understand him as emphasizing that a description of the self that leaves out reference to the drives is incomplete or misleading.

Nietzsche's claim that the self is a relation of drives is also designed to make a second point. If you want to describe a particular human being, one of the most informative things that you can do is to describe what the person cares about, what she values, and what she characteristically strives toward. Indeed, one thing that we mean when we ask for descriptions of a "true self" is a description of what the person really values, really wants, really cares about. And, as the previous chapters explained, the person's cares, values, and characteristic actions are largely determined by her drives; the default position is valuing the ends of one's drives (cf. D 109). In light of this, when Nietzsche says that the self is a structure of drives, we should understand him as claiming that the self is characterized by what it values.

Passages such as BGE 268 make this perfectly explicit:

Which group of sensations [*Empfindungen*] is aroused, expresses itself, and issues commands in a soul most quickly, is decisive for the whole order of rank of its values and ultimately determines its table of goods. The values of a human being betray something of the *structure* of his soul and where it finds its conditions of life, its true need. (BGE 268)

The structure of the self is given by the structure of the individual's values. Just so, Nietzsche claims that in order to determine what one truly is, one must ask:

What until now have you truly loved, what has raised up your soul, what ruled it and at the same time made it happy? Line up these objects of reverence before you, and perhaps by what they are and by their sequence, they will yield you a law, the fundamental law of your true self [*deines eigentlichen Selbst*]. Compare these objects, see how one completes, enlarges, exceeds, transforms the other, how they form a ladder on which you have so far climbed up toward yourself. (UM III.1)

And a bit later:

Your true nature [*dein wahres Wesen*] does not lie hidden deep inside you, but immeasurably high above you, or at least above that which you customarily consider to be your ego. (UM III.1; cf. HH I.624)

These passages are alike in identifying the self with the agent's values. So rather than making the reductive claim that the self is equivalent to the drives, Nietzsche means to emphasize that the self is defined by its values—values that are largely set by the drives.

8.1.2 *Descriptive and Aspirational Senses of Selfhood*

If we want to understand a self, then, we need to specify what it values. And these values will be set largely by the individual's drives. But this raises a question. "Self" is often used as a descriptive term, picking out the characteristic features of human individuals. Thus, Harry Frankfurt claims that human selves are characterized by their capacity to entertain higher-order attitudes (Frankfurt 1988); Gary Watson argues that the distinguishing feature of selfhood is the capacity for evaluation (Watson 1975); and David Hume adopts a reductive account, arguing that selves are merely bundles of attitudes (Hume 1739/1969, Book I, Part IV, Section VI). In each case, the account of selfhood purports to pick out a feature that is standardly present in adult human beings.

At first glance, Nietzsche's account might seem analogous; he might seem to be claiming that what is characteristic of human beings is just having drive-induced values. But this isn't quite right. Nietzsche usually treats selfhood as an honorific or aspirational term. This is clear in some of the passages quoted above. For example, UM III.1 distinguishes the "true self" from what the person already is. Elsewhere in that work, Nietzsche writes, "Be yourself! All you are now doing, thinking, desiring is not you yourself" (UM III.1). GS 335 claims that we "want to become those we are," and urges people to "create themselves." HH II.366 claims that if you "*will* a self you shall *become* a self." We are not selves simply in virtue of being human. Indeed, Nietzsche claims that most of us *lack* selves: "we absolutely should not assume that many human beings are 'people' [*Personen*]," he tells us (KSA 12:10[59]). Selfhood is something that must be attained, rather than something that we automatically possess.¹

How should we make sense of this? What distinguishes the ordinary person from the genuine self? Here, there are several tempting readings. We've seen that one strand in Nietzsche's thought emphasizes that most of us are fragmented and disunified;

¹ See Gemes (2001).

accordingly, we might treat selfhood as attained when we unify ourselves. I addressed a version of this topic in Chapter Seven, arguing that Nietzsche develops a notion of unity that is meant to pick out cases in which the agent is active in the production of her action (this type of unity obtains when, were one to know more about the etiology of one's action, one's attitude toward this action would be stable). This is a condition of agential activity rather than selfhood, and it is entirely compatible with being a fragmented, incoherent self who flits from activity to activity. Of course, Nietzsche also does speak of unity in other senses; for example, he often praises agents whose drives are unified in common projects. That isn't a plausible candidate for agential activity, but perhaps it is for selfhood. Or perhaps Nietzsche has some alternative account of selfhood, such as being a great individual, or freeing oneself from the grip of culture, or selecting one's own values. I'll explore these possibilities in turn below.

8.2 The Unified Self

8.2.1 *Nietzschean Selfhood as Having Unified Drives or Unified Activities*

A number of commentators read Nietzsche's remarks on unity as designed to present a picture of selfhood. Ken Gemes notes that "For Nietzsche, unified selves, what he takes to be genuine persons, are rare achievements" (Gemes 2009b, 49). And he suggests that being a genuine person involves having drives that are unified in the pursuit of common objects: "For Nietzsche, in certain rare cases, drives, rather than being split off, are harnessed into a centered, unified whole" (2009b, 47). In particular, this occurs through "the redirection of those drives away from their initial, primary goal toward a secondary goal that is more in line with the master drive" (Gemes 2009b, 47). So Gemes claims that we count as selves when a master drive sublimates other drives, thereby eliminating conflict and directing us toward a single project.²

We find similar views in other interpreters. As noted in Chapter Seven, Reginster (2003) treats unity as obtaining when one drive masters, rather than tyrannizes or represses, the other drives. Richardson (1996, 33) discusses the way in which drives can rule other drives by redirecting them toward ends more consonant with the dominant drive. Walter Kaufmann presents the higher man as someone who "has overcome his animal nature, organized the chaos of his passions, sublimated his impulses, and

² More recently, Gemes has qualified this position to some extent. Although he still tends to identify being a self with being unified, Gemes allows that Nietzsche's texts can be read as distinguishing selfhood and unity: "In this essay we have more or less explicitly identified a range of notions found in Nietzsche's texts—becoming what you are, having a self, becoming a genuine individual, being a person—with that of achieving a unity of drives under the sublimating hegemony of a master drive. Alternatively, one might wish to argue that such unity is for Nietzsche only a prerequisite of becoming a self and/or becoming what you are, and/or becoming a genuine individual, and/or being a person; and to provide an account that differentiates between these ideas. Or, inversely, one might claim that being a person, and/or having a self, is a precondition for establishing unity. Nietzsche's texts are arguably not precise or detailed enough to provide a definitive basis for strict distinctions of this kind" (Gemes and Le Patourel, forthcoming).

given style to his character" (1974, 316). Although these accounts of unity vary in their details, they agree in presenting selfhood as attained through unification of the drives or affects.

Another group of commentators focuses less on unity between internal constituents, such as drives, and more on outward manifestations of unity. For example, Alexander Nehamas argues that selfhood requires a form of narrative unity among the events in one's life: "Nietzsche's model for the world, for objects and for people turns out to be the literary text and its components" (1985, 90). In particular, we should strive to manifest "the perfect unity and freedom that are primarily possessed by perfect literary characters" (1985, 195). We do this by fashioning our actions, character, and goals "into a controlled and coherent whole" (1985, 227). Thomas Hurka eschews this narrative model, but agrees in seeing the relevant kind of unity as obtaining not so much between drives as between the agent's goals or projects (2007).

Although we might quibble with the details of these views, there's no denying that Nietzsche praises *some forms* of unity between drives, aspects of one's character, or, more generally, the activities composing one's life. GS 290 is illustrative: there, Nietzsche valorizes those who practice the "great and rare art" of "giv[ing] style to one's character" by reshaping themselves in light of some "artistic plan." Analogously, EH II.9 speaks of directing one's life with an "organizing 'idea.'" BGE 200 laments the fact that modern individuals have "drives and value standards that fight each other and rarely give each other any rest." Nietzsche has Zarathustra complain that he "always discovers the same thing: fragments and limbs and dreadful chances—but no men! . . . I walk among men as among fragments of the future: of the future which I scan. And it is my art and aim, to compose into one and bring together what is fragment and riddle and dreadful accident" (Z II.21). And we saw in Chapter Seven that Nietzsche writes:

The multitude and disgregation of impulses and the lack of any systematic order among them results in a "weak will"; their coordination under a single predominant impulse results in a "strong" will: in the first case it is the oscillation and lack of gravity; in the later, the precision and clarity of direction. (KSA 13:14[219])

All of these passages emphasize the fragmentation and disunity of modern individuals. Many of the passages treat this as a distinctively modern predicament, something brought about by the nature of modern values or modern social structures; the ancients, with somewhat more harmonious drives, and with social structures that allow direct expression of more drives, did not experience the same degree of inner division.³

In light of this evidence, we might conclude that Nietzschean selfhood requires one or more of these forms of unity. Absent that unity, we fail to qualify as selves. In short, we might interpret Nietzsche as adopting the position that was famously defended by Hölderlin:

³ See, for example, Gemes (2006) on the way in which modern persons tend to be fragmented.

There are two ideals of our being: one is a condition of the greatest simplicity, where our needs are in accordance with one another, within our powers, and with everything we are related to, *just through the organization of nature*, without any action on our part. The other is a condition of the highest cultivation, where this accord would come about between infinitely diversified and strengthened needs and powers, *through the organization that we are able to give ourselves*. (Hyperion Fragment)

Hölderlin treats unity of motivation, powers, and activities as an ideal, and distinguishes a natural unity in simplicity from a cultivated unity in complexity. On a standard reading, this is basically Nietzsche's view as well. The self, for Nietzsche, would be the one who cultivates some kind of unity in her mental economy or her actions. Moreover, the ancients may not have needed to do this, or may not have needed to do so in quite the same way; their unity came more naturally to them.

8.2.2 *Problems with the Interpretation*

Given the frequency and intensity with which Nietzsche praises these forms of unity, it's easy to see why readers are tempted to identify selfhood with unity. I don't want to deny that Nietzsche sometimes uses the term "self" in that way; nor do I wish to deny that he values some forms of unity. However, I'll argue that unity is not his central concern. He is typically interested in unity not because he thinks it has any intrinsic import, but because of what it tends to produce or exemplify. There are three ways in which we can see this.

First, it is important that when Nietzsche praises these forms of unity, he typically praises them *for their effects*. That is, he seems to treat unity of drives or activities as having instrumental, rather than intrinsic, value. The oft-quoted GS 290 is illustrative: the title of that passage is "One thing is needful," and it begins with a discussion of artistic unity. This misleads some commentators into thinking that the one needful thing is achieving unity. But Nietzsche explicitly tells us what the one needful thing is: toward the end of the passage, he writes, "For one thing is needful: that a human being should *attain* satisfaction with himself." So he is presenting unity of activities as a means to self-satisfaction, rather than an end in itself.⁴ The passages on unity from *Zarathustra* make a similar point: the fragmentation of modern individuals is problematic in that it renders them incapable of affirming their lives. Analogously, passages such as BGE 200 claim that disunity *weakens* us and thereby renders us unable to affirm ourselves: "such human beings of late cultures and refracted lights will on average be weaker human beings: their most profound desire is that the war they *are* should come to an end" (BGE 200). Importantly, in this passage Nietzsche immediately goes

⁴ I read GS 290 as claiming that artistic unity is only instrumentally valuable, as a means to attaining satisfaction with oneself. But there is a second way of reading this passage: we could interpret Nietzsche as *identifying* artistic unity with a certain kind of self-satisfaction, so that *what it is* to manifest the relevant type of unity just is to be satisfied with oneself. If we read Nietzsche in this way, then the type of unity he is discussing in GS 290 turns out to be intimately related to, and perhaps identical to, the kind of unity I analyzed in Chapter Seven.

on to praise *disunity* “when the opposition and war in such a nature have the effect of one more charm and incentive of life.” For then it can give rise to:

Those magical, incomprehensible, and unfathomable ones..., those enigmatic men predestined for victory and seduction, whose most beautiful expression is found in Alcibiades and Caesar... and among artists perhaps Leonardo da Vinci. They appear in precisely the same ages when that weaker type with its desire for rest comes to the fore: both types belong together and owe their origin to the same causes. (BGE 200)

All of these passages suggest that Nietzsche values unity of drives or activities only insofar as it enables affirmation. If this is right, then unity of drives or activities has only instrumental value. This makes it implausible to analyze Nietzschean selfhood solely in terms of unity. (Assuming, that is, that selfhood has more than just instrumental value.)⁵

We can strengthen this conclusion by considering a second point: stepping back from the details of Nietzsche’s texts, it is hard to see how unity of drives or activities could have any intrinsic value. Suppose I live an erratic, haphazard, disorganized life, flitting from project to project, giving no thought to any grand artistic plans or all-encompassing projects, but nonetheless bearing a wholly affirmative attitude toward life. What is supposed to be wrong with that? Why would this life be improved if I pressed it into some coherent narrative? I see no answers to these questions. Or, approaching this from a different direction, would the disunified individuals whom Nietzsche mentions—Alcibiades, Caesar, da Vinci—somehow be improved if they had more uniform, coherent drives or more tightly integrated projects? This seems implausible. There seems to be no principled reason for insisting that selfhood demands these forms of unity.

Of course, Nietzsche could just be using “self” as an honorific. That is, he clearly does value some of the forms of unity discussed above, and it’s possible that he simply uses the term “self” to pick out these valued forms. But this would be of limited interest: nothing would then turn on whether a person is labeled a self or not. We could dispense with the talk of selfhood and simply focus on unity of drives or affects,

⁵ These considerations aren’t decisive: although the passages discussed above do seem to treat unity as merely a condition for something valuable, there are other passages that can be read as suggesting that unity is valuable in itself. This emerges most clearly in Nietzsche’s discussions of culture: he tends to treat cultures with a unity of style as superior to cultures that contain disparate and competing styles (see, for example, UM I.1–2, BGE 224, CW 7, TI IX.11, A 58–9). But I would make two points. First, even these passages on cultural unity tend to be qualified: for example, whereas UM I.1 praises cultures that exhibit “unity of artistic style in the expressions of the life of a people,” the next section claims that even “inferior and degenerate” cultures can exhibit unity (UM I.2). So it seems that cultural unity is, at best, one component of a flourishing culture. Second, the passages that most strongly suggest that unity is valued in itself tend to identify unity not with mere coordination of drives or projects, but instead with a unity of *taste or style* (see, again, the passages cited in this paragraph). What Nietzsche is interested in here is a certain *aesthetic* value that he describes as unity: the individual or culture is an aesthetically appealing whole. If unity is being understood in terms of aesthetic criteria, then it is the aesthetic criteria—rather than unity itself—that are doing the work. I have no quarrel with the idea that Nietzsche sometimes evaluates persons and cultures on aesthetic grounds. There’s no doubt that he does. But I want to investigate a rather different question: whether this kind of unity is criterial for selfhood. I see no indication that it is.

instead. This is the third and most pressing problem: if “self” just refers to unity of drives or affects, then there is no philosophical motivation for speaking of selves. The term “self” wouldn’t identify or explain anything that isn’t already explained in other terms.

8.3 The Individual’s Relationship to Society

The unity accounts, though diverse in their details, share two common features. First, they are purely formal: they specify selfhood in terms of some type of coherence or order, without reference to the substantive features of the self (what she values, what she desires, what her projects are, how she relates to others, etc.). Second, they treat relationships between the person and her social context as adventitious: it may be that the unified person’s activities involve others and put her in relation to the values and customs of her society, but this would only be accidental.

These two features are responsible, I think, for the account’s problems. When we look more closely at Nietzsche’s treatments of persons, we find that he is deeply concerned both with the particular, substantive activities that they are engaged in and with the way in which they relate to their social context.

To draw out these points, I want to look, briefly, at Nietzsche’s discussions of “great” or “higher” individuals. In his texts, Nietzsche repeatedly presents certain historical individuals as instructive models. Goethe, Napoleon, and Beethoven are the most frequent objects of positive attention; Schopenhauer, Plato, and Wagner get a more mixed reaction.⁶ So we might determine how Nietzsche envisions selfhood by seeing what is distinctive about these exemplars.

Why does Nietzsche label individuals certain individuals—most frequently, Goethe, Beethoven, and Napoleon—“higher” or “great”?⁷ For one thing, these higher individuals are committed to immensely difficult goals (BGE 260). Moreover, they make considerable progress in attaining these goals, but never rest content with their current level of accomplishment; instead, they constantly seek new challenges (cf. Z II.12). But that is hardly unique to them: there are many individuals who lead lives full of struggle and accomplishment without coming in for much praise from Nietzsche. So we need to find a feature that sets the great individuals apart.

⁶ For example, he writes that Napoleon is “the first man, and the man of the greatest tasks and advanced views, of modernity” (KSA 11:37[9]); Schopenhauer is “the last German worthy of consideration” (TI IX.21); Goethe “created himself” (TI IX.49) and is “the last German for whom I feel any reverence” (TI IX.51); Plato is “a first rate decadent” (TI X.2); Wagner is the “most instructive case” for a “*diagnosis of the modern soul*” (CW Epilogue). In the eKGW edition, Goethe is mentioned 633 times, Beethoven 229, and Napoleon 220. The passages on Goethe are uniformly positive; the ones on Beethoven and Napoleon are mostly positive. Other figures are discussed more frequently than these—Wagner is mentioned over 2,500 times, Schopenhauer over 1,000 times, Plato over 600—but the passages on these figures are far more mixed.

⁷ I’ve discussed these points in greater depth in Katsafanas (2013, 232–4).

There is an obvious candidate: what's common to the great individuals Nietzsche mentions is that they have profound social and cultural impacts. Goethe's literature influenced a whole generation.⁸ Beethoven played a transformative role in musical history and served, for many nineteenth-century thinkers, as a paradigm of artistic genius (BGE 245). Napoleon had an obvious impact on European society and culture. (Nietzsche makes similar points about the impact of Schopenhauer, Goethe, and Wagner.)⁹ All of these individuals manifest "a long, terrible will" that is able to "cast its goals millennia hence" (BGE 208). A person whose projects are more private affairs, with no broader impact on culture, would not be a candidate for Nietzschean greatness. So one criterion for greatness is *the extent of one's impact on society*.

Indeed, one way of seeing this point is by considering *Zarathustra*, which Nietzsche regarded as his greatest work. This book is in large part a description of Zarathustra's attempts to influence others. The book opens with Zarathustra rejecting a solitary existence and seeking an audience (Z Preface). Though his attempts initially misfire, and though he progresses from seeking to reach the masses to cultivating a small group of followers, his goal is always to spread a new set of values or a new ideal across society. The idea of an isolated sage, an isolated great individual, is rejected throughout: Zarathustra is presented as going into solitude only when he believes that doing so is a way of correcting personal failings (Z II.22 and Z III.1).

It's clear, then, that Nietzsche's great individuals have large cultural impacts. But is this sufficient for greatness? Hurka takes a version of that line, identifying Nietzschean greatness with a merely formal condition: the extent to which the individual's activity encompasses a large number of persons and objects, as well as the extent to which all of the agent's goals are involved in this overall activity (Hurka 2007, 23–4). However, this is too simple: not just any form of extensive impact is sufficient for greatness. So, for

⁸ Nietzsche writes, "*Goethe*—not a German event, but a European one: a magnificent attempt to overcome the eighteenth century by a return to nature, by an *ascent* to the naturalness of the Renaissance—a kind of self-overcoming on the part of that century. He bore its strongest instincts within himself: the sensibility, the idolatry of nature, the anti-historic, the idealistic, the unreal and revolutionary (the latter being merely a form of the unreal).... What he wanted was *totality*; he fought the mutual extraneousness of reason, senses, feeling, and will (preached with the most abhorrent scholasticism by *Kant*, Goethe's antipode); he disciplined himself to wholeness, he *created* himself.... he said Yes to everything that was related to him... and he had no greater experience than that *ens realissimum* called Napoleon. Goethe conceived a human being who would be strong, highly educated, skillful in all bodily matters, self-controlled, reverent toward himself, and who might dare to afford the whole range and wealth of being natural, being strong enough for such freedom; the man of tolerance, not from weakness but from strength, because he knows how to use to his advantage even that from which the average nature would perish; the man from whom there is no longer anything that is forbidden—unless it be *weakness*, whether called vice or virtue. Such a spirit who has *become free* stands amid the cosmos with a joyous and trusting fatalism, in the *faith* that only the particular is loathsome, and that all is redeemed and affirmed in the whole—he *does not negate anymore*" (TI IX.49).

⁹ "Schopenhauer, the last German worthy of consideration (who represents a *European* event like Goethe, like Hegel, like Heinrich Heine, and not merely a local event, a 'national' one), is for a psychologist a first-rate case: namely, as a maliciously ingenious attempt to adduce in favor of a nihilistic total depreciation of life precisely the counter-instances, the great self-affirmations of the 'will to life,' life's forms of exuberance" (TI IX.21). For remarks on Wagner's influence, see BT, CW, and NCW. For Plato, see for example TI.II.

example, Luther and St. Paul have tremendous cultural impacts, unifying great multitudes of individuals in the service of one overarching goal, but Nietzsche condemns them because of the particular shifts that they bring about.¹⁰ So shifting culture is a necessary but not sufficient condition for greatness.¹¹

What more is required? The content of the agent's project seems to matter. In particular, Nietzsche presents all of his great individuals as manifesting an affirmative attitude toward life: these higher individuals exemplify "the ideal of the most high-spirited, alive, and world-affirming human being who has not only come to terms and learned to get along with whatever was and is, but who wants to have what was and is repeated into all eternity" (BGE 56; cf. BGE 287, EH II.2, TI IX.49). Indeed, this is what seems to separate the influential figures whom Nietzsche counts as great—figures such as Goethe, Napoleon, and Beethoven—from influential figures he condemns or, at least, is not inclined to valorize: individuals like Schopenhauer, Plato, and Wagner. Nietzschean greatness requires an affirmative attitude toward life.

Moreover, notice that Nietzsche's great individuals shift culture *by instituting or promoting new values*. In Nietzsche's words, these great individuals are all "untimely" or independent from their age, in the sense that they are innovators in artistic values, political values, and so forth. As Nietzsche puts it, "the concept of greatness entails being noble, wanting to be by oneself, being able to be different, standing alone and having to live independently" (BGE 212; cf. BGE 211, GS 98). Thus, Nietzsche claims that Napoleon represents "the problem of the noble ideal as such made flesh" (GM I.16); he is "the ideal of antiquity incarnate" (GM I.16). (Notice that Nietzsche focuses more on Napoleon's ideal than his military and political exploits.) Goethe is "a magnificent attempt to overcome the eighteenth century by a return to nature, by an *ascent* to the naturalness of the Renaissance—a kind of self-overcoming on the part of that century" (TI IX.49). And Beethoven is characterized as a key figure in the shift to Romantic values (BGE 245). All of these higher individuals free themselves from conventional values and institute new ideals. (Plato, Schopenhauer, and Wagner do the same, but promote values with contents that Nietzsche regards as non-affirmative.)

¹⁰ Nietzsche claims that Luther blocked an incipient revaluation of values: "The Germans did Europe out of the harvest, the meaning, of the last great age, the age of the Renaissance, at a moment when a higher order of values, the noble ones, those that say Yes to life, those that guarantee the future, had triumphed at the seat of the opposite values, those of decline—even in the very instincts of those who were sitting there. Luther, this calamity of a monk, restored the church and, what is a thousand times worse, Christianity, at the very moment when it was vanquished." (CW 2). Nietzsche presents Paul as largely responsible for the "slave revolt"—the transition from classical to Christian values—that he describes in the *Genealogy*. For example, Nietzsche writes that Paul "guessed how one could use the little sectarian Christian movement apart from Judaism to kindle a 'world fire'; how with the symbol of 'God on the cross' one could unite all who lay at the bottom, all who were secretly rebellious...into a tremendous power. Christianity as a formula with which to outbid the subterranean cults of all kinds...and to unite them: in this insight lies the genius of Paul. His instinct was so sure in this that he took the ideas with which these chandala religions fascinated, and, with ruthless violence, he put them into the mouth of the 'Savior,' whom he had invented" (A 58; see also A 41–57).

¹¹ Hurka briefly considers this type of objection in Hurka (2007, 26–7).

In particular, these individuals provide what Nietzsche sometimes terms “higher values.”¹² Higher values are the goals that can “inspire faith” (KSA 12: 9[35]): they “promote the life of the species, *by promoting the faith in life*. ‘Life is worth living; every one of them shouts, ‘there is something to life, there is something behind life, beneath it; beware!’” In short, these higher values provide life with a meaning: “life *ought to be* loved, *because—!*” (GS 1). They provide “an image of man”—an image of what we should aspire to be (UM III.4). Nietzsche gives some examples:

There are three images of man which our modern age has set up one after the other and which will no doubt long inspire mortals to a transfiguration of their own lives: they are the man of Rousseau, the man of Goethe, and finally the man of Schopenhauer. (UM III.4)

Nietzsche tells us that Rousseau’s ideal is the return to nature: the natural is identified with the good, the social with corruption. Goethe idealizes the “contemplative man in the grand style,” the person who takes delight in discovery. And Schopenhauer valorizes the ascetic seeker after truth: “The Schopenhauerean man voluntarily takes upon himself the suffering involved in being truthful, and this suffering serves to destroy his own willfulness and to prepare that complete overturning and conversion of his being, which it is the real meaning of life to lead up to” (UM III.4). These examples could be multiplied.

In sum, Nietzsche’s great individuals aren’t just those who have extensive impacts on society. Rather, they are those who have extensive impacts *with particular, life-affirmative contents and through a particular method*: promotion of new values.¹³

I’ve focused on great individuals because they help us to draw out three important features of Nietzschean selfhood. First, his account is not merely formal: the content of the agent’s actions and commitments matters. Second, these individuals are not solitary, but deeply engaged with their cultures. Third, their projects involve or constitute revaluations of values.

While the great individuals instantiate these features to high degrees, I’m not aware of any passages in which Nietzsche suggests that terms such as *self*, *person*, and so forth apply *only* to these great individuals. So when he urges us to become who we are or to realize genuine selfhood, he isn’t demanding that we strive for this kind of evaluative impact on culture. And for good reason: this form of greatness is not something to which very many individuals can coherently aspire. It requires the tremendous good luck to be born at a moment in which society is ripe for redirection and the individual

¹² For a discussion of the nature and import of higher values, as well as their relation to Nietzsche’s concerns about nihilism, see Katsafanas (forthcoming b).

¹³ It’s interesting to wonder why this particular form of cultural impact is singled out. After all, inventors, rulers, physicians, industrialists, and so forth also have large impacts on society—arguably much larger impacts than Goethe and Beethoven—but are of no interest to Nietzsche. He is interested in those who create new ideals or provide new values. This may seem strange—why are shifts in values so much more important than shifts in industry or economy?—but it is Nietzsche’s view. (In a way, we might see Nietzsche as the inverse of Marx: whereas Marx focuses solely on the economic factors driving social change, Nietzsche focuses solely on the philosophical and religious ones.)

is capable of providing that redirection. A very few individuals may find themselves, through sheer luck, living at these historical junctures and in possession of the capacities to shift values; but this cannot happen with any regularity, and can only be the privilege of a few individuals across the millennia.

Being a great individual is *sufficient* for attaining selfhood; but it does not seem *necessary*. However, with these features of greatness in view, we can specify a more restricted criterion for selfhood. Few will have immensely influential projects. But it's easier to achieve something else: revaluing values, and in particular striving to commit oneself to affirmative values. That is, the Nietzschean self might be the person who, rather than passively accepting the values of her society, instead fashions her own values, and fashions them in a way that constitutes affirmation rather than negation. In other words, we can separate the criteria for greatness from the criteria for individuality. I suggest that what makes the great individuals *individuals* is that they revalue values; what makes them *great* is that their revaluations have extensive influence on their cultures. I'll defend this reading below.

8.4 Nietzschean Selfhood as Authenticity

I'll argue that Nietzsche's aspirational notion of selfhood is this: a person qualifies as a self to the extent that she revalues her values. I'll review passages to this effect below. But first I need to introduce a distinction.

Broadly speaking, there are two traditions that identify the attainment of selfhood with some form of independence in valuing:

- The genuine self is the authentic self: she clears away the distorting influences of culture and cultivates her own uniqueness. Call this the Romantic view.
- The genuine self is the autonomous self: she critically assesses values and adopts only those values that she freely endorses. Call this the Kantian view.

I'll explain these pictures in the following sections. Both have resonances with Nietzsche's writings: he sometimes seems to be endorsing a version of each, and accordingly interpreters often attribute one of these two views to him. I think the truth is more complex. Each of these views is responsive to themes in Nietzsche's texts. However, Nietzsche is actually striving to articulate a view of selfhood that incorporates the best elements of the Romantic and Kantian views, while avoiding their excesses. I'll articulate this view below. The remainder of this chapter will address the Romantic view; I'll turn to the Kantian view in Chapter Nine.

8.4.1 Richardson's Account

Perhaps the best place to begin is John Richardson's account in *Nietzsche's New Darwinism*, which exhibits elements of both the Romantic and the Kantian views. Richardson's core idea is that we are shaped by three forces: natural selection, social

selection, and individual or self-selection. Our drives are set by natural selection; our customs and values are set by social selection; and the free individual chooses his values by self-selection.

Richardson's notion of social selection requires some explanation. Social selection involves a capacity to remember and copy the characteristic behaviors of other members of one's social group (Richardson 2004, 81–2). As Richardson puts it:

This capacity sets up a new replicative mechanism, in which habits get copied and thereby multiply in a nongenetic way. And a new selective logic will work on these new replicators: they will be selected for success at copying and dispersing themselves as social customs. How well a habit enhances genetic fitness may be one factor affecting how well it spreads, but there can be many others. (Richardson 2004, 82)

In short, we have a capacity to copy habits or customs, and this capacity favors habits or customs that others will also adopt. So easily imitated habits and customs are propagated within society, whereas ones that are more unique or difficult to copy tend to die out.

Richardson claims that these customs tend to conflict with the drives. Social selection works "*against* natural selection. It works to modify behaviors designed for the organism's survival and reproduction, and to re-aim them toward goals serving a different overall end. So taught behaviors oppose or counter inherited behaviors" (2004, 84). Because drives resist this pressure, "customs must constantly work to oppose and suppress inherited drives" (2004, 84).

Summarizing these points, Richardson writes:

On the one hand our bodies and drives have been adapted for an animal, physical health—for fitness to further the reproducing line. This involves a deep egoism. But our "spirit"—our habits of memory, consciousness, and language—has been adapted by selection over a different field of replicators, for a different and in fact conflicting end. Habits are selected to bind us into society, and the most effective means, under both the ethic of custom and morality, are habits that attack and undermine our drives and their natural "healthy selfishness." Hence we find ourselves—and this is the sum of Nietzsche's dual diagnosis—"the sick animal," "human, all-too-human": our two deep projects frustrate and interfere with one another, and subject us to a pervasive suffering. (Richardson 2004, 120)

So social selection instills in us a set of behavioral dispositions that are in conflict with our drives. And this is where the third form of selection, self-selection, becomes important: "The self-creation or freedom that Nietzsche means lies in bringing the selective process into oneself. It lies in taking over, oneself, the selective role, so that one 'creates' or 'gives oneself' values" (2004, 96):

The key to becoming myself is to select my values, i.e., the goals of the dispositions that—in making my behavior—specify "who I am." It is to make myself the cause of these decisive bodily aims. Initially and as a matter of course, these aims are set by selection in the species' or society's past: how my body aims was selected—"became"—before I was born, within the natural and social processes that formed my drives and habits. To become myself is to make my

values *during* my life, as well as by and for that life. I revise these drives and habits, selecting them to serve my individual will. (2004, 96)

We do this by achieving insight into the origins and true functions of our values and incorporating this insight, striving to shift our motives and habits. In so doing, we can aim to eliminate the conflict between natural and social selection (2004, 120–4). As Richardson puts it, “The argument is, in a way, dialectical and Hegelian. The conflict between these opposing principles (natural and social selection, with their distinct values) is resolved in a new synthesis” (2004, 124).

There’s much to admire in this account. I think it’s true that Nietzsche recognizes a form of social selection that operates with just the logic Richardson describes. And, as I’ll argue in Chapter Nine, Richardson’s self-selection is a good summary of what the sovereign or autonomous individual does (with the caveat that, on my reading, the sovereign individual assesses values in terms of power). So, in broad outlines, Richardson’s triple distinction between forms of selection seems exactly right.

However, I think Richardson errs in an important respect. I’ll critique two different aspects of these views: (1) the Romantic association of the true self with the drives and society with corruption; (2) relatedly, the Kantian idea that we ought to aim at a form of autonomy that is based solely on standards issuing from the independent, detached self. These two ideas are clearest in Richardson’s account, but one or more of them are present in many influential readings of Nietzsche.

8.4.2 *Nietzsche’s Early Acceptance of the View*

A crucial component of Richardson’s interpretation is the idea that social selection instills behavioral dispositions that are contrary to an individual’s own interest. In his early works, Nietzsche does seem to accept a version of this claim. Indeed, he seems to accept a quite extreme version of it: the Romantic claim that society and custom tear the person away from some underlying uniqueness. The third *Untimely Meditation* is the best example. There, Nietzsche writes that “every person is a unique miracle” (UM III.1), for “each of us bears a productive uniqueness within him as the core of his being; and when he becomes aware of it, there appears around him a strange penumbra which is the mark of his singularity” (UM III.3). We are each unique, precious, singular selves, but we are tempted away from this uniqueness by our immersion in culture: “the person who does not wish to belong to the mass needs only to cease taking himself easily; let him follow his conscience, which calls to him: ‘Be your self! All you are now doing, thinking, desiring, is not you yourself’” (UM III.1). In endeavoring to be himself, the individual comes into conflict with his social context:

for he is contending against those aspects of his age that prevent him from being great, which means, in his case, being free and entirely himself. From which it follows that his hostility is at bottom directed against that which, though he finds it in himself, is not truly himself: against the indecent compounding and confusing of things eternally incompatible, against the soldering of time-bound things on to his own untimeliness; and in the end the supposed child of his time proves to

be only its stepchild. Thus Schopenhauer strove from his early youth against that false, idle, and unworthy mother, his age, and by as it were expelling her from him, he healed and purified his being and rediscovered himself in the health and purity native to him. (UM III.3)

Here, Nietzsche wholeheartedly endorses the idea that the genuine self rips himself free of culture, strips away all that is extraneous and foreign, and discovers, buried deep down inside, some uncorrupted, natural core, some “productive uniqueness.” He ceases “thinking and acting as part of a herd,” instead “joyously being himself” (UM III.5).

This is a paradigmatically Romantic picture of individuality: the true self is located by stripping away the effects of culture.¹⁴ It is a rather simplistic picture, open to a number of objections. For example: why assume that the self is pre-social, that I have deep commitments that can be understood as prior to and concealed by my culture? Do we find some productive uniqueness, some font of originality, in the slaves of the *Genealogy* or the fragmented moderns of whom, in a notebook entry from 1887, Nietzsche writes that “Many are multiple people, and most are none. . . . the ‘person’ is a relatively isolated fact” (KSA 12[10]:59)? It’s clear that as soon as we move past his early works, Nietzsche answers these kinds of questions negatively. For example, BGE 266 attributes the following quotation to Goethe: “one can only have true respect for one who does not *seek* himself.” A passage from *Zarathustra*, quoted above in Section 8.1, makes the same point: “it is my art and aim, *to compose into one and bring together* what is fragment and riddle and dreadful accident” (Z II.21, emphasis added).

After the early works, the passages on selfhood routinely treat selfhood as something to be created rather than some preexisting unity to be discovered.¹⁵ So, if we confine our attention to Nietzsche’s more mature views, Ken Gemes seems entirely correct in writing that “Nietzsche rejects the romantic notion that we were at one time whole but now are sundered. Indeed, Nietzsche radically reverses the temporality of unity; where romantics typically locate unity in the past . . . Nietzsche posits unity as a possible future achievement” (Gemes 2001, 355).

But what causes this shift in Nietzsche’s view? I think he detects two problems with the Romantic view: there is no such thing as a pre-social drive, and human nature is malleable.

¹⁴ Consider, for example, how Adrian del Caro explains Romanticism: the Romantics engage in a project of “seeking out the essential, natural vestiges of humanity least corrupted by culture and prior to the ascendancy of reason” (2013, 109). They preach “an elevation of the embodied virtues, e.g., feeling, the passions, spontaneity, physical activity, and creativity associated with the unconscious or the irrational. The ‘good’ characters in their respective novels are those who live close to nature . . . these natural types are capable of staving off the corrupting and enslaving tendencies of culture and society, both of which are too addicted to reason, rote learning, quantified knowledge, and subjugation of nature” (2013, 109). Del Caro points out that we can understand BT as endorsing something like that picture, with the Dionysian standing in for the Romantic notion of nature.

¹⁵ In fact, Nietzsche’s commitment to this Romantic view is never wholehearted; there are already some tensions emerging in UM III. For example, Nietzsche writes, “Your true nature [*dein wahres Wesen*] does not lie hidden deep inside you, but immeasurably high above you, or at least above that which you customarily consider to be your ego [*dein Ich*]” (UM III.1).

8.4.3 First Problem: There Is No Such Thing as a Pre-Social Drive

Consider Richardson's suggestion that naturally selected drives aim at individual preservation while socially selected customs and habits aim at their own preservation (Richardson 2004, 82–4). I think Nietzsche's view is more complex. Many of our drives are inherently *social*. To pick the most obvious example, Nietzsche mentions a “drive toward sociality” (WS 70), but also a drive to distinguish oneself, a drive to religion, a drive to sex, to property, and to play.¹⁶ In HH I.49, he refers to an “unegoistic drive” that manifest itself as “good-naturedness, friendliness, politeness of the heart” as well as “pity, compassion, and self-sacrifice.” He frequently speaks of a drive to knowledge, claiming that it is fostered by “higher education” (HH I.254). HC and other texts (such as WS 226) focus on the drive toward competition, which is given determinate forms by social contexts.

There are two points to make about these drives. First, while Richardson certainly acknowledges the existence of these sorts of behavioral dispositions, he treats them as socially selected habits and customs, rather than drives (see, for example, Richardson 2004, 84 and 120). I cannot discern any textual basis for doing so. Nietzsche calls these inherently social behaviors drives (*Triebe*), and he seems to treat them as coeval with the more egoistic drives.

Second, these drives aren't inhibited by the propagation of behavioral patterns in society; on the contrary, they cannot function without such patterns. Drives are always manifest against some cultural backdrop, and the particular ways in which they motivate us depend on cultural standards. The sex drive, the knowledge drive, the drive to express domination, and so on, all have their conditions set in part by the social context in which they operate. For what counts as pursuing sex, knowledge, domination, and so on, is in part determined by what the society recognizes as sex, knowledge, domination, and so on. Particular customs might inhibit these drives in particular contexts, but custom as such is a condition for them, rather than an obstacle to them.¹⁷

To amplify this point, consider the fact that certain cultures don't exhibit widespread conflict between drives and social customs. It's a distinctive feature of Judeo-Christian morality or, as Nietzsche sometimes calls it, “anti-natural” morality that it does. Nietzsche does not think there was systematic conflict between drives and customs in the Homeric Greeks (cf. HC and GM I). And he wonders whether it might be possible to reshape society so that the bad conscience is experienced when the agent acts on “all the unnatural inclinations, all those aspirations to the beyond, to that which runs counter to sense,

¹⁶ For these drives, see: D 113, 143; “Truth and Lie” Section 2; KSA 7:3[64]; KSA 7:16[13]; UM II:6, HH 281, AOM 211, WS 31, TI IX.39.

¹⁷ We might be misled by the fact that some parts of the *Genealogy* sound like *Civilization and its Discontents*, in that they suggest that society necessarily stifles various drives (see esp. GM II.16). However, a careful reading of GM II.16 and similar passages reveals that Nietzsche is *not* suggesting that *all* drives are stifled by socialization. Particular aggressive drives are redirected by socialization, but it would be a mistake to conclude that this is common to all drives. I discuss this point in more detail in Katsafanas 2005.

instinct, nature, animal, in short all ideals hitherto, which are one and all hostile to life and ideals that slander the world” (GM II.24).

So it's a mistake to think that Nietzsche endorses the Romantic idea that the self is something that is necessarily apart from and in conflict with culture. The fact that there are many conflicts between *particular* customs and *particular* drives does not entail that we should accept a *custom/drive* dichotomy, or that we should see the drives as things that can be understood as having pre-social aims.

8.4.4 *Second Problem: Human Nature Is Malleable*

Another problem with the Romantic view is that, in urging us to return to some pre-social nature, it ignores the ways in which human nature is historically fluid. Consider Nietzsche's complaint about figures such as Rousseau: Nietzsche interprets Rousseau as calling for a “return to nature” (TI IX.1, IX.48), and thereby assuming there is some determinate, pre-social essence that has been corrupted by culture.¹⁸ We can see the same error more clearly articulated in the sentimentalists. Hume, for example, writes:

It is universally acknowledged that there is a great uniformity among the actions of men, in all nations and ages, and that human nature remains still the same, in its principles and operations.... Mankind are so much the same, in all times and places, that history informs us of nothing new or strange in this particular. Its chief use is only to discover the constant and universal principles of human nature, by showing men in all varieties of circumstances and situations, and furnishing us with materials from which we may form our observations and become acquainted with the regular springs of human action and behavior. (*Enquiry on Human Understanding*, Of Liberty and Necessity, Part 1).

Nietzsche rejects this view. Human nature is fluid: it is modified by our values, our customs, our habits. He makes this point as early as *Human, All too Human*:

Family failing of philosophers.—All philosophers have the common failing of starting out from man as he is now and thinking they can reach their goal through an analysis of him. They involuntarily think of “man” as an *aeterna veritas*, as something that remains constant in the midst of all flux, as a sure measure of things. Everything the philosopher has declared about man is, however, at bottom no more than a testimony as to the man of a very limited period of time. Lack of historical sense is the family failing of all philosophers; many, without being aware of it, even take the most recent manifestation of man, such as has arisen under the impress of certain religions, even certain political events, as the fixed form from which one has to start out. (HH I.2)

¹⁸ Nietzsche's characterizes Rousseau's “image of man” as follows: “man cries in his distress to ‘holy nature’ and suddenly feels that it is as distant from him as any Epicurean god. His prayers do not reach it, so deeply is he sunk in the chaos of unnaturalness. Scornfully he throws from him all the gaudy finery which only a short time before had seemed to him to constitute his essential humanity, his arts and sciences, the advantages of a refined life; he beats with his fists against the walls in whose shadow he has so degenerated, and demands light, sun, forest and mountain. And when he cries: ‘Only nature is good, only the natural is human’” (UM III.4).

And he echoes the point in *Twilight*:

You ask me which of the philosophers' traits are really idiosyncrasies? For example, their lack of historical sense, their hatred of the very idea of becoming, their Egypticism? They think that they show their respect for a subject when they de-historicize it, *sub specie aeterni*—when they turn it into a mummy. All that philosophers have handled for thousands of years have been concept mummies; nothing real escaped their grasp alive. (TI III.1)

In short: “there has never been a natural mankind” (KSA 12:10[53]/WLN 185). Philosophers have erred in taking human nature as static when it is in fact fluid.

Consider how this bears on drives. Drives and the values they induce take very different shapes in different social contexts. To choose one of the simplest examples, Nietzsche thinks we have drives that dispose us to vent cruelty on others (GM II.5–6). He emphasizes this point at length, writing that in the ancient world it is openly acknowledged that “to make someone suffer is pleasure in its highest form” (GM II.6). He continues, “cruelty is part of the festive joy of the ancients, and, indeed, is an ingredient in nearly every pleasure they have” (GM II.6). An example: “Not so long ago it was unthinkable to have a royal wedding or full-scale festival for the people without executions, tortures or perhaps an auto-da-fe, similarly, no noble household was without creatures on whom people could discharge their malice and cruel taunts with impunity” (GM II.6). He sums up this attitude as follows: “To see suffering does you good, to make suffer, better still” (GM II.6).

Of course, this has changed. Consider one example of the change: “remember *Don Quixote* . . . today we read the whole of *Don Quixote* with a bitter taste in the mouth, it is almost an ordeal, which would make us seem very strange and incomprehensible to the author and his contemporaries—they read it with a clear conscience as the funniest of books, it made them nearly laugh themselves to death” (GM II.6). Although we, too, have drives toward cruelty (GM II.11), we do not engage in these festivals of cruelty and, indeed, find them reprehensible. Nietzsche famously suggests that our drive toward cruelty instead finds *internal* expression: the drives expressed as “hostility, cruelty, joy in persecuting, in attacking, in change, in destruction” are turned against their possessor, generating “an animal soul turned against itself, taking sides against itself” (GM II.16). Thus:

The man who, from lack of external enemies and resistances and forcibly confined to the oppressive narrowness and punctiliousness of custom, impatiently lacerated, persecuted, gnawed at, assaulted, and maltreated himself; this animal that rubbed itself raw against the bars of its cage as one tried to “tame” it; this deprived creature, racked with homesickness for the wild, who had to turn himself into an adventure, a torture chamber, an uncertain and dangerous wilderness—this fool, this yearning and desperate prisoner became the inventor of the “bad conscience.” (GM II.16)

Cut off from opportunities to vent our aggression externally, we turn against ourselves. As Nietzsche puts it a bit later:

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

The ancient and the modern persons both have drives toward cruelty, but the drives spark very different activities and values depending on the social context in which the individual finds himself. The reason for this is obviously not just that the ancient and the modern consciously and explicitly decide to value different things. Consider just how alien the ancient perspective is to us: consider how reprehensible we would find these displays of cruelty. This is deeper than mere conscious judgment. It is a modification of our affects, our ways of seeing the world.

These sorts of reflections give the lie, I think, to Richardson’s idea that there is a straightforward conflict between the egoistic ends set by our drives and the unegoistic ends set by social customs. The drive toward cruelty in the ancients took one form; in us, it takes another. The activities of the ancients would not better suit our drives; it’s not as if, were we to begin torturing and barbarizing others, we would suddenly find our drives enjoying complete fulfillment. As Nietzsche puts it in *Daybreak*, our drives have been, and are continually being, “transformed by moral judgments” (D 38).

The general point, here, is that the ways in which drives manifest themselves and induce values is dependent on external, non-individual factors. For the drives are gradually shaped and transformed by their setting.

An important aside: we should be wary of going too far in the other direction. The fact that human nature changes over time does not entail that it is *completely* mutable. Thus, in BGE 230 Nietzsche enjoins us to clear away misinterpretations and discover “the terrible basic text of *homo natura* . . . For to translate man back into nature; to master the many vain and fanciful interpretations and secondary meanings which have hitherto been scribbled and daubed over that eternal basic text *homo natura* . . . that may be a strange and extravagant task but it is a *task*.” To give a simple example, the fact that we’ve long interpreted ourselves as possessing libertarian freedom does not make it the case that we actually have libertarian freedom. The fact that we’ve long interpreted our intellects as enjoying independence from our motives does not make it the case that our intellects actually do enjoy this independence.

There is variability in our drives and affects, and hence in our values. And there is some variation in our capacities: for example, Nietzsche thinks that we once lacked, but now possess, the capacity to promise, to bind ourselves, to recognize commitments (GM II). We’ve acquired an intellectual conscience (albeit a weak one) and a drive toward truth (GM III).

So human nature is historically variable not in the sense that broad structural features—such as whether we have free will or independent egos—change over time, but

instead in the sense that our deepest commitments and highest values, our needs and aspirations, shift in profound ways over time.

(This, incidentally, is why a Nietzschean philosophical psychology is compatible with his claim that human nature is historically fluid. The topics investigated in the previous chapters—the relation of conscious to unconscious, the nature of drives, and the relationship between drives, values, and will—all of these concern largely fixed aspects of human nature. The topic we are investigating in this chapter do not.)

8.5 Conclusion: From Authenticity to Autonomy

I suggest that these two problems—the rejection of pre-social drives and the recognition that human nature is malleable—lead Nietzsche to reject the Romantic view. If we tried to strip away the accretions of culture, we wouldn't find a true, unblemished self; we would find nothing at all. The self has to be created, not discovered. So genuine selfhood cannot be identical to detached, pre-social individuality.

Nonetheless, Nietzsche does seem to associate the ideal of genuine selfhood with those who counteract some of the tendencies of their culture and create new values. So, even if he rejects the idea of pre-social selves, might he endorse the idea that genuine selfhood is to be attained by engaging with and critiquing social customs, norms, and values? Might he accept something like Richardson's notion of self-selection, which involves reflecting on our values, affects, and commitments, and freely assessing them?

There are some initial reasons for skepticism. If Nietzsche really were presenting the free individual as the form of life toward which we should all aspire, then he would be recapitulating eighteenth- and nineteenth-century orthodoxy. Consider Kant:

Enlightenment is man's emergence from his self-imposed immaturity [Unmündigkeit]. Immaturity is the inability to use one's own understanding without another's guidance. This immaturity is self-imposed if its cause lies not in lack of understanding but in indecision and lack of courage to use one's own mind without another's guidance. *Dare to know! (Sapere aude.)* "Have the courage to use your own understanding," is therefore the motto of the enlightenment... This enlightenment requires nothing but *freedom*. (Kant 1784/1959)

The enlightened individual is the one who critically assesses his own convictions and values, daring to use his own reason, breaking free of his habitual acceptance of the convictions and values of his society. Likewise, John Stuart Mill writes that "the human faculties of perception, judgment, discriminative feeling, mental activity, and even moral preference are exercised only in making a choice. He who does anything because it is the custom makes no choice... He who lets the world, or his own portion of it, choose his plan of life for him has no need of any other faculty than the ape-like one of imitation" (1859/1978, 56). In Kant and Mill, then, the ideal is the life of autonomous self-direction. Rather than passively accepting the values of our societies, we use our own understanding to determine our own values. Could Nietzsche really accept this Kantian and Millian orthodoxy?

Certainly Nietzsche has reservations. He doubts that most individuals will be capable of attaining freedom (GS 18, GS 98, GS 347), and he presents it as a dangerous aspiration for most of us (EH II.9). His account of motivation's pervasive effects on reflection entails that most attempts to assess values will be driven by unnoticed background affects and commitments; that, in short, these allegedly autonomous assessments will be a farce, driven by the very values the agent purports to suspend. But we could view these points as nothing more than tweaks to the Kantian/Millian picture: all but the most hardened Kantians would agree that *if* striving for autonomy were as ruinous for the average man as Nietzsche describes, then we would have reason to avoid it. And Kantians have no problem acknowledging that the attainment of autonomy is rare and difficult. So what we need to ask is whether this kind of Kantian/Millian freedom is really what Nietzsche is associating with selfhood, as Richardson suggests.

I'll argue that it is not. The problem is not just that Nietzsche tends to critique what's taken as philosophical common sense. In addition, he directly calls into question some aspects of the traditional picture of autonomous self-direction. His problem with the Millian account is obvious: he rejects the idea, essential to Mill's account, that the autonomous agent can direct his conduct in light of objective values (for Mill, pleasure and pain). His objections to Kant are subtler. Recall BGE 21's critique of the idea that we are *causa sui*, self-caused causes. Nietzsche condemns the desire "to bear entire and ultimate responsibility for one's actions oneself, and to absolve God, the world, ancestors, chance, and society," for this desire "involves nothing less than to be precisely this *causa sui* and, with more than Münchhausen's audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness" (BGE 21). While Nietzsche is here focused on actions, the point extends more broadly: part of what he rejects is the idea of self as pure interiority, capable of being understood in isolation from its social and historical context. The Kantian picture assumes that the self can generate substantive normative standards in a solely internal fashion, drawing on nothing but its own understanding. Instead, Nietzsche thinks we should understand the self as porous, as shaped and largely defined by social and historical forces (after all, the drives are defined in these ways, and so too are what the individual takes as final values). Or, put a bit differently: just as it would be a profound mistake to ignore the way in which the unconscious is intertwined with the conscious, so too it would be a mistake to examine the individual without seeing the way in which social factors meld her.

So BGE 21's critique of the will as *causa sui* isn't *just* a critique of libertarian accounts of freedom; it suggests that the whole way of conceiving of the self as something independent of its context is problematic. Kant envisions the self generating normative standards out of the bare idea of freedom or rationality, but I'll argue that for Nietzsche, the self always begins with determinate values and affects set by its social context. These can't be understood, in the Kantian fashion, as external to the will; they are constituents of the will.

Nonetheless, in Chapter Nine I will suggest that Nietzsche does have a conception of freedom. It is a conception that rejects the Kantian ideal of a self-defining, autonomous agent. But it does involve a knowing assessment of the values in which we find ourselves immersed. And attaining it does, indeed, vouchsafe a certain form of genuine selfhood.

The Nietzschean Self

*Moral Psychology, Agency, and
the Unconscious*

Paul Katsafanas

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Paul Katsafanas 2016

The moral rights of the author have been asserted

First Edition published in 2016

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2015948399

ISBN 978-0-19-873710-0

Printed in Great Britain by
Clays Ltd, St Ives plc

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.