

## Value, Affect, Drive

Paul Katsafanas

pkatsa@bu.edu

Boston University

Penultimate draft. Forthcoming in Peter Kail and Manuel Dries (eds.), *Nietzsche on Mind and Nature*. Oxford: Oxford University Press.

Nietzsche's writings display a pervasive concern with values. He diagnoses the pathologies inherent in our commitment to certain values; he traces the history of our values; he argues that some values have horrific psychic costs; he warns us that our current values may be "the danger or dangers"; and he urges us to create new values. Left out of these warnings, critiques, and exhortations, unfortunately, is a clear explanation of what exactly a value *is*. What is it for a person to value something?

The answer to that question might seem straightforward. We typically view values as manifest in the agent's reflective judgments. To value X is, in part, to make judgments of the form "X is valuable," to guide one's practical deliberation with principles such as "X ought to be promoted," and so forth. Valuing X is simply a matter of being committed to certain reflective judgments and acting in accordance with them.

This would be a natural view of values, but Nietzsche seems to deny it. For although Nietzsche does devote some attention to reflectively espoused evaluative judgments, he more often focuses upon valuations inherent in pre-reflective psychological phenomena: our drives. He frequently claims both that drives *include* evaluations and that drives *explain* reflective evaluative judgments.

Consider a typical passage. Nietzsche writes "a drive without some kind of knowing evaluation... does not exist in man" (HH I.32). Remarks of this kind suggest that Nietzsche identifies valuing X with having a drive toward X. But if *that* is Nietzsche's view, we straightaway encounter philosophical difficulties. After all, Nietzsche claims that we are largely ignorant of our drives:

However far a man may go in self-knowledge, nothing however can be more incomplete than his image of the totality of drives which constitute his being. He can

scarcely name even the cruder ones: their number and strength, their ebb and flood, their play and counterplay among one another, and above all the laws of their nutriment remain wholly unknown to him. (D 119)

If we are ignorant of our drives, and each drive includes a valuation, then it follows that we are ignorant of many of our own values. That is not how philosophers tend to think of values: as I pointed out above, we typically regard values as manifest in reflective judgments.

To make matters worse, Nietzsche also claims that the agent's conscious judgments about what is valuable are to be *explained* in terms of drives. "Our moral judgments and evaluations," Nietzsche tells us, "are only images and fantasies based on a physiological process unknown to us (D 119; cf. BGE 19). On examination, this claim is rather mysterious. Manifestly, the agent's reflective judgments about value can diverge from the aims of his drives (this is, after all, one of Nietzsche's complaints about modern morality). For example, an agent might judge that pursuit of power is disvaluable, while being strongly driven toward the pursuit of power.

In short, although Nietzsche clearly wants to associate values with pre-reflective drives, and to explain reflective judgments about value in terms of these drives, it is unclear how he can do so. The agent's reflectively endorsed ends, principles, and commitments—what we would think of as the agent's *values*—seem not only distinct from, but often in conflict with, the agent's drives.

For these reasons, it is unclear how we should understand Nietzsche's concept of value. This essay will attempt to dispel the puzzles surrounding that concept by offering an account of the nature of Nietzschean value. Section One begins by considering *drives*. I explain what a drive is and consider John Richardson's claim that Nietzsche identifies values with the goals of drives. I argue that while Richardson's view captures important aspects of Nietzsche's view, it stands in need of supplementation. Section Two examines the connection between *affects* and values. Peter Poellner, Maudemarie Clark, and David Dudrick have argued that Nietzschean values can be identified with affects that are regarded as justified. I argue that this view, as it stands, faces significant difficulties. However, I suggest a modification of this view that enables it to overcome the objections. According to the view that I defend, an agent values X iff (i) the agent has a drive-induced affective orientation toward X and (ii) the agent does not disapprove of this affective orientation. I

contend that this view makes sense of Nietzsche's remarks about value and overcomes the difficulties inherent in the competing interpretations.

If this view is to succeed, though, it must also account for Nietzsche's claim that drives *explain* the agent's reflective judgments concerning what is valuable. Section Three investigates this aspect of Nietzsche's view. I argue that drives generate thoughts about justification, thereby strongly inclining agents to regard pursuit of the drive's end as valuable. Section Four considers a potential objection to this view, an objection that has recently been raised by Sebastian Gardner: if I come to recognize that my reflective valuations are drive-induced, won't this undermine my commitment to those values? How could I simultaneously accept a valuation and regard it as a product of unreflective drives? I show that we can answer this objection by appealing to the recalcitrance of drive-induced affective orientations.

## 1. Drives include evaluations

In a number of passages, Nietzsche claims that drives either *include* or *constitute* evaluations of their goals:

A drive without some kind of knowing evaluation of the worth of its objective, does not exist in man. (HH I.32)

From each of our basic drives there is a different perspectival valuation [*perspektivische Abschätzung*] of all events and experiences. (KSA 12, 1[58])

In all willing is esteeming... (KSA 11, 25[433])

willing: = willing an end. End contains a valuation. (KSA 10, 24[15])

These passages claim that merely *having a drive toward X* entails *having an evaluation of X*. To make sense of this idea, Section 1.1 begins by asking what a drive is. With that account in place, Section 1.2 considers Richardson's claim that Nietzsche identifies values with the ends of drives. Section 1.3 argues that there are potential problems with Richardson's

interpretation. Examining these problems will disclose criteria of adequacy for an account of Nietzschean value.

### 1.1 What a drive is

In Katsafanas (forthcoming), I argue that drives have four key features: (i) they are dispositions that generate affective orientations; (ii) they admit an aim/object distinction; (iii) they dispose agents to seek their aims, rather than their objects; and (iv) they are constant. I will explain these points in turn.

First, a drive is a disposition that generates an affective orientation. Drives manifest themselves by structuring the agent's perceptions, affects, and reflective thought. They determine which features of the agent's environment will be salient; they determine the way in which the agent will conceptualize aspects of the environment; and they sometimes alter the course of the agent's reflective thought. For example, when the sex drive is active, the agent will find potential sexual partners salient; he will tend to view others in terms of their status as potential sexual partners, and will view objects and events in his environment (a bar, a conversation) as instruments or obstacles toward sexual activity; and his thoughts will tend to circle around topics related to sexuality. I will use the term "affective orientation" to refer to this system of affectively charged perceptions and thoughts.

Second, drives admit an aim/object distinction. The aim of a drive is its characteristic form of activity, in terms of which it is distinguished from other drives. For example, the aim of the sex drive is sexual activity; the aim of the aggressive drive is aggressive activity. In order to express their characteristic form of activity, drives need to find particular objects: to express aggressive activity, I need to find someone or something to be aggressive toward. These temporary goals, upon which the drive vents its activity, are the drive's *objects*. They are highly variable and may bear only the most tenuous relationships to one another. Thus, the object of the aggressive drive might be another individual, but it might just as well be a punching bag, a video game, or a bout of aggressive driving.

This brings us to the third point: what the drive seeks is expression of its *aim*. The aggressive drive doesn't dispose the agent to attain any particular goal, but simply to engage in aggressive activity (its aim). In order to express this form of activity, the drive needs to find objects. But the attainment of these objects does not satisfy the drive, for these objects

are simply chance occasions for expression. Thus, after expressing itself on one object, the drive will seek another. While the aim is constant, the objects can vary continuously. In short, the drive motivates the agent to express a certain form of activity (aggressive activity, sexual activity etc.), thereby fulfilling its aim; in expressing this activity, the agent has temporary objects (aggressive activity in sport, romantic interaction with a particular individual, etc.).

Finally, drives are constant. Here, a contrast with desires is helpful. In typical cases, a desire dissipates once its goal is attained. If I have a headache and desire to take an aspirin, the desire will be extinguished once I take the aspirin. Drives are different. The sex drive may be temporarily sated by a sexual encounter, but it will soon arise again. The aggressive drive may be put into abeyance by a good game of basketball, but it will recur soon enough. There is no object the attainment of which will eliminate the drive.<sup>1</sup>

## **1.2 Richardson on the relationship between drives and values**

With this account of drives at hand, let's consider how drives might relate to values. Given Nietzsche's constant association of values with drives, it is very tempting to read Nietzsche as identifying values with the aims of drives. John Richardson offers an influential defense of this interpretation. Below, I argue that although Richardson captures something important about Nietzsche's view, the interpretation needs refinement in order to avoid two problems: it is forced to identify mere cravings with valuations, and it renders the connection between reflective and unreflective values mysterious. Seeing the ways in which Richardson's interpretation encounters these problems will reveal why the powerful temptation to read Nietzsche as identifying values with the aims of drives must be resisted.

Richardson begins by claiming that Nietzsche identifies values with the ends of drives:

He takes the role of valuing... away from a central ego-will-mind, and disperses it among a multitude of drives... each drive is a valuing. (Richardson 2004: 74)

---

<sup>1</sup> See Katsafanas (forthcoming) for a full explanation and defense of these claims, as well as an analysis of the way in which drives differ from ordinary desires.

Spelling out this view in more detail, Richardson writes that “a drive’s value is simply the outcome it tends plastically (and responsively) toward” (Richardson 2004: 75).

This requires a bit of explanation. Drives are “plastic,” “responsive” dispositions in the following sense:

A drive is a **plastic** disposition to [its outcome], inasmuch as it tends to produce different behaviors in different circumstances, in such a way that the same outcome is reached, by different routes, in all of them... Such plasticity depends on a capacity to ‘respond’ to circumstances... in some minimal way. (Richardson 2004: 75).

In other words, the plasticity of a drive consists in its ability to generate different behaviors in different circumstances, such that the drive’s aim will be achieved in all of them. We can illustrate this point with a simple example: satisfying the hunger drive will require hunting in one circumstance, picking up a fork in another, going in to a shop in still a third, and so on. Or, to use one of Nietzsche’s examples: the aggressive drive may find expression in harming others, but it may also change directions, finding expression in self-inflicted cruelty. Thus,

if one route [to its goal] is blocked, it shifts to another. The disposition bifurcates (trifurcates, and so on) but then reconverges: it is a tendency to respond differently in different contexts, in such a way that the same result ensues. (Richardson 2004: 28)

The drive motivates a variety of behaviors that may appear distinct, but which serve a common goal.

Generating these variable behaviors requires a “responsiveness and self-adjustment to environmental conditions” on the part of the drive (Richardson 2004: 29). If I see that I cannot gratify my hunger drive by pulling something out of the refrigerator, I walk over to the local restaurant instead; if I see that I cannot express cruelty on others, I adapt, expressing it on myself. The drive’s ‘value’ is to be identified with this end, rather than with

the temporary object. Using my aim/object terminology, the drive's value is the *aim*, rather than the object.<sup>2</sup>

We might wonder how drive's aims—which Richardson is identifying with values—relate to the agent's conscious thoughts about value. In response to this kind of question, Richardson argues that we should distinguish two kinds of value: reflective, conscious values and unreflective, bodily values. He writes,

valuing need not—and principally does not—occur in a conscious act... We suppose that “our values” are those we put into *language* and *consciousness*... But according to Nietzsche... the really effective or influential values are not those conscious ones, but values we have, as it were, through the plant or animal in us. Values are built into our bodies, and their conscious and linguistic expression is something quite secondary. (Richardson 2004: 73-74)

Here, Richardson suggests that consciously espoused values can be regarded as secondary, whereas values “built into our bodies” can be regarded as primary. Elsewhere, he explains:

although Nietzsche cares not at all about how an amoeba or a cat or a chimp values, he thinks that an ‘animal’ kind of valuing operates in us all the time, in the ‘part’ of us we call our body. I’ll refer to this lower level of valuing in us as body-values; these are the ends in us, that we share with other living things. Nietzsche speaks of this valuing whenever he speaks of our drives. But of course this isn’t what we usually call our ‘values’, which are the values we’re aware of and put into words, the principles we refer to in order to steer our actions. They are our ethics or morality. By contrast with body-values, I’ll call these agent-values. These are conscious and linguistic, the rules or ideals we formulate and then refer to in ‘moments of decision’.

(Richardson forthcoming)

---

<sup>2</sup> Richardson's claims about plasticity are closely related to what I have described as the aim/object distinction. However, I view the aim/object distinction as more fundamental than, and explanatory of, the drive's plasticity. In particular, to say that a drive is a plastic disposition is to say that it motivates different behaviors in different circumstances, where all of these behaviors are paths to the same goal. The aim/object distinction explains this feature: the drive is plastic precisely because it selects a variety of objects that provide opportunities for expressing its aim.

Body-values are the values that are present in our drives; agent-values are our consciously espoused valuations, such as “compassion is good.”

Let’s provisionally accept Richardson’s distinction between unreflective, bodily values and reflective, agential values. We can put the point this way:

Unreflective values = the aims of our drives

Reflective values = our consciously espoused evaluative judgments

I’ve discussed Richardson’s view at some length because I think it represents an extremely tempting way to read Nietzsche. Nietzsche’s phrasings invite us to distinguish reflective and unreflective values, to treat the latter as the *real* values, and to identify these real values with the aims of our drives. Tempting as this interpretation is, though, I will argue that it must be refined.

### 1.3 Potential problems for this interpretation

The first question to ask about unreflective values is why, exactly, the aims of our drives should be labeled *values*. Given Richardson’s account of plasticity and responsiveness, we can see why the label is tempting. If I have a plastic, responsive disposition toward some goal, a variety of apparently heterogeneous behaviors will have an underlying unity—they will be done in the service of that goal. Moreover, I will be disposed to modify my behavior—sometimes in very dramatic ways—in order to achieve that goal. This seems to conform to some of our ordinary thoughts about values. Values are typically regarded as things that structure large swathes of our actions, and things that have a kind of tenacity, requiring us to modify our behaviors in changing circumstances to achieve them.

However, a problem arises when we consider the fact that not every goal toward which we are disposed counts as a value. Cases of addictions provide good examples: imagine an alcoholic struggling to overcome his addiction to alcohol, which he regards as reprehensible. This agent will be strongly disposed toward consuming alcohol, in much the way that an ordinary agent will be strongly disposed, by his reproductive drive, toward sexual activity. We could describe the alcoholic urge as a plastic, responsive disposition: after all, the urge will manifest itself in a variety of ways (desire to go to a bar, desire to go to a liquor



store, desire to go to a party, etc.) and it will motivate new behaviors when it finds one path blocked (going to a party rather than the bar which turned out to be closed, etc.). Nonetheless, the addict hardly seems to *value* alcohol. If this is right, then we cannot identify values with plastic, responsive dispositions.

Of course, we could say that this agent reflectively disvalues alcohol and unreflectively values alcohol, but this simply postpones the question of why the disposition toward drinking alcohol should count as a *value*. It is unclear what would justify that claim.

This brings us to a second difficulty. Suppose we do accept the distinction between reflective and unreflective values. As defined above, it might appear that reflective values are completely unrelated to unreflective values. After all, my consciously espoused evaluative judgments might be wholly distinct from, or indeed even opposed to, the aims of my drives. For example, an ascetic might consciously espouse a valuation such as “sexual activity is evil,” while having sexual activity as the aim of one of his drives. But Nietzsche doesn’t accept the view that unreflective values and reflective values are wholly distinct. Indeed, as I mentioned in the introduction, he often suggests that unreflective values—drives—explain reflective values.

In sum, while interpreting values as the aims of drives is tempting, the interpretation faces certain difficulties. First, we need to understand why drive-induced aims should count as values. Second, we need to see how drive-induced aims relate to the agent’s consciously espoused evaluative judgments. These are criteria of adequacy on a successful interpretation of Nietzsche on drives.

## 2. Affects and values

So far, we have focused on the passages in which Nietzsche asserts a connection between values and drives. However, there are also a number of passages in which Nietzsche draws tight connections between values and *affects*. Consider a few examples:

Moralities are also merely a sign language of the affects. (BGE187)

It is clear that moral feelings are transmitted in this way: children observe in adults inclinations for and aversions to certain actions and, as born apes, *imitate* these

inclinations and aversions; in later life they find themselves full of these acquired and well-exercised affects and consider it only decent to try to account for and justify them... (D 34)

You still carry around the valuations of things that originate in the passions and loves of former centuries! (GS 57)

In these passages, Nietzsche claims that there is some relationship between affects and values. Unfortunately, it is not clear what sort of relationship he has in mind. BGE 187 suggests that evaluations *express* affects. GS 57 claims that valuations “originate in” affects, which seems to imply a distinction between values and affects. D 34, on the other hand, seems to claim that evaluations are *constituted* by affects that are regarded as justified. In the following sections, I will clarify Nietzsche’s position by asking exactly what the relationship between values and affects is supposed to be. Peter Poellner, Maudemarie Clark, and David Dudrick have developed related views of Nietzsche on value. I think these views come very close to the correct interpretation of Nietzsche, so I begin by reconstructing them below. I then point out a crucial shortcoming of the views, and propose a new interpretation that overcomes it. By seeing that and how these encounter difficulties, we will be in a position to develop an adequate account of Nietzsche on value.

## **2.1 The Poellner/Clark/Dudrick view**

In an insightful article, Poellner argues that Nietzschean values are “co-constituted” by affects and their intentional objects. Poellner begins by distinguishing two uncontroversial features of values. First, a value judgment involves “a commitment to an attitude of favouring or disfavouring towards some features of the judgment’s object” (Poellner 2007: 230). In other words, part of what it is to value (say) generosity is to be committed to an attitude of favoring certain aspects of generosity.

Second, value judgments “appear as objective” in the sense that “evaluative commitments, in order to be even minimally intelligible, must be guided or constrained in some way” (Poellner 2007: 231). In other words, value judgments appear to be constrained

by features of their objects. In valuing generosity, I take this attitude to be warranted by features of generosity.

Poellner clarifies this latter point as follows:

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

For example, consider a perfectly ordinary evaluative experience. I might experience a stunning seaside vista as valuable. In so doing I experience an attraction to that vista, and I see this attraction as justified by the vista's character. In particular, I see it as merited by the vista's beauty. Here we have both features highlighted above: an attitude of favoring together with the thought that this attitude is not arbitrary, but is merited by the object in question.

We can view these as conditions of adequacy on an account of value: anything that is a plausible candidate for a value must have these two features. Poellner argues that *affective experiences* meet the conditions (Poellner 2007: 234-48). In particular, Poellner argues that some emotions can be regarded as perceptions of value. According to Poellner, there are certain "perceptual emotions" that present their intentional objects as having evaluative properties. For example, an emotion of awe might represent itself as being caused by the beauty of a vista. Or, an emotion of admiration directed at generous behavior might represent itself as being caused by the goodness of generosity. In this respect, the emotion exhibits its intentional object as causal in virtue of its evaluative properties.

If this is right, then the two conditions listed above are met. The first condition is met because the perceptual emotions clearly involve an attitude of favoring. The second condition is met because the perceptual emotions appear as objective, in the sense that they appear to be not merely caused, but also merited, by their objects.

Thus, Poellner argues that certain affects can be interpreted as perceptions of first-personal attractiveness, and that this perception of X as attracting, together with the thought

that this attraction is merited by X's features, constitutes a valuing of X. Put simply, to value X is to have a positive attitude toward X together with a thought that this positive attitude is warranted.

Clark and Dudrick (2007) offer a similar interpretation. They focus on the way in which values are created, and, in particular, on how values *spread* from one person to another. They write that we create values by “inducing people to posit things as valuable,” and we do this by “by inducing in people new affective responses to things” (Clark and Dudrick 2007: 212). Thus, part of what it is to value X is to have a certain kind of affective response to X. However, a further condition is needed:

One must induce people not only to develop dispositions to act and react to things in certain ways, but also to take these ways of acting and feeling to be justified, to be supported by reasons. Nietzsche does this in his own writings by offering new *interpretations* of the dispositions he induces in people (to which he seduces them), ones designed to induce people to regard those dispositions as justified. (Clark and Dudrick 2007: 212)

To instill a new value, one must cause the agent develop an affect together with the sense that this affect is rationally warranted. In short,

To value something—as opposed to merely desiring it—is to accept that there is *reason* to take certain actions and attitudes towards it (for instance, to approve or disapprove of it, and to do or refrain from doing it)... (Clark and Dudrick 2007: 213)

In these passages, Clark and Dudrick emphasize that valuing X is not just having a positive affective orientation toward X. In addition, one must regard this positive affective orientation as *justified*.

Both of these interpretations agree, then, that to value something is to have an affective orientation toward it; but merely having an affective orientation isn't enough to constitute valuing, because it is an essential feature of the concept *value* that in order to value something one must take the affective orientation as merited or justified. So, we might summarize the Poellner/Clark/Dudrick view as follows:

Valuing X = (1) Having a positive affect toward X, and (2) regarding this affect as merited or justified.

This interpretation of values seems consistent with the passages cited above.

Consider, again, the passage from *Daybreak*:

It is clear that moral feelings are transmitted this way: children observe in adults inclinations and aversions to certain actions and, as born apes, *imitate* these inclinations and aversions; in later life they find themselves full of these acquired and well-exercised affects and consider it only decent to try to account for and justify them... (D 34)

Here, Nietzsche does seem to assert that valuations are “feelings” or “inclinations” that are regarded as justified. If this is his view, values would indeed originate in affects, as GS 57 asserts. Moreover, values could be regarded as a “sign-language” (BGE 187) of the affects in the sense that each valuation includes an affect. Thus, this interpretation seems supported by textual evidence. Might it succeed?

## **2.2 Refinements: affects must be drive-induced**

The Poellner/Clark/Dudrick view, like Richardson’s view, captures something true and important about Nietzsche’s account. Nonetheless, the view does face a problem. It seems implausible to identify values with affects-regarded-as-justified, for a simple reason: I can have passing, fleeting affects that I regard as warranted but that do not count as values. Suppose that once in my life, I have an attraction to skydiving. I have affective experiences of approval toward the thought of skydiving, and I regard this affect as justified by the belief that the experience will be exhilarating. Nonetheless, the fancy passes, and I never give it a second thought. It would hardly be right to say that I *value* skydiving, even at the moment when I have that affective response. A “value” held only for a moment is no value at all.

Consequently, values can be identified only with a subset of the affects regarded as justified. And, in fact, there is an easy way to pick out the relevant subset: we can combine the Richardson view with the Poellner/Clark/Dudrick view. Suppose that we identify values

with the affects *that are induced by drives*. Above, I pointed out that drives operate, in part, by generating affects that incline the agent to pursue the drive's aim (by presenting certain courses of action in an appealing light). These drive-induced affects will be more stable and abiding than ordinary affects, and they will be centered on a common aim. For example, consider an agent under the influence of the sex drive. This drive will lead the agent to have persistent experiences of attraction toward opportunities for sexual activity. Given that this attraction will be relatively constant, and will be present in a wide range of behaviors sharing the common attribute of providing objects for the drive's aim, it makes sense to speak of the agent as valuing sexual activity.

This identification of values with drive-induced affects is further supported by the fact that drives do not induce discrete, specific emotions toward their aims and objects. Rather, each drive induces an affective orientation. The sex drive doesn't simply generate a brute inclination toward copulation: it generates a range of attractions, emotions such as love and joy, propensities toward activities that render one attractive to potential sexual partners, and so forth. This system of interrelated affects and dispositions is what I have above labeled an "affective orientation." It differs from discrete affects, such as having a temporary attraction to skydiving, in that it involves a host of interconnected, mutually reinforcing affects and behavioral dispositions.

It seems more plausible to identify valuations with these affective orientations than with discrete affects. For again, "value" typically refers to something that exerts a structuring effect on broad swathes of the individual's life. So, let's provisionally identify values with drive-induced affective orientations that are regarded as justified.

### **2.3 A problem for the view**

We have made progress in uncovering Nietzsche's account of value. However, a lacuna remains: we need to ask what it is to regard an affective orientation as justified. Are we to imagine an agent explicitly entertaining the thought "my affective orientation is justified"? Or is something subtler involved?

Clark and Dudrick suggest that an affect counts as a value only if the agent takes there to be reasons for having the affect. As they put it, to value something "is to take it that there are reasons (i.e., it is rational) to feel and act towards it in certain ways" (Clark and

Dudrick 2007: 214). However, this interpretation is bound to strike readers of Nietzsche as excessively intellectualistic. There are two ways of making this point.

First, and most obviously, Nietzsche often speaks of agents who have values but give no thought whatsoever to the rational warrant of these values. Consider a typical passage:

The fettered spirit takes up his position, not for reasons, but out of habit; he is a Christian, for example, not because he has knowledge of the various religions and has chosen between them... he encountered Christianity... and adopted it without reasons, as a man born in a wine-producing country becomes a wine drinker. Ask a fettered spirit for his reasons against bigamy, for example, and you will learn whether his holy zeal for monogamy is based on reasons or on habit. (HH I.226)

The Christian's disvaluation of bigamy and valuation of monogamy is based on habit, rather than reason. But the Christian nonetheless values monogamy. So it seems that reflecting on reasons is not a necessary condition for having a value.

The same point is made elsewhere:

Your judgment, 'this is right' has a prehistory in your drives, inclinations, aversions, experiences, and what you have failed to experience; you have to ask, 'how did it emerge there?' and then also, 'what is really compelling me to listen to it'. (GS 335)

This passage suggests that an agent can have an evaluative judgment (e.g. "this is right") without inquiring into its warrant. After all, if valuing something *required* considering warrant, there would be no need for Nietzsche to urge us to inquire into warrant.

So the first problem with the Clark/Dudrick proposal is that many agents have values but give no thought to their justificatory status. This brings us to a second problem: Nietzsche often speaks of unnoticed values, values that agents are not even aware of having. Consider the following passage:

What compels one to this, however, this unconditional will to truth, is the belief in the ascetic ideal itself, even if as its unconscious imperative—do not deceive yourself about this,—it is the belief in a metaphysical value, a value in itself of truth as it is

established and guaranteed by that ideal alone (it stands and falls with that ideal).  
(GM III.24)

In this passage, Nietzsche suggests that many scientists and philosophers value asceticism: in pursuing truth, they unreflectively commit themselves to treating asceticism as an ideal. But these thinkers hardly entertain thoughts such as “my affective orientation toward asceticism is justified.” Indeed, these thinkers do not even realize that they have a positive affective orientation toward asceticism. Judging from this passage, it would seem that we can value something without knowing that we have an affective orientation toward it. Of course, we cannot regard an affective orientation as justified or rationally warranted if we are unaware even of *having* the orientation.

So there are two ways in which the Poellner/Clark/Dudrick interpretation seems excessively intellectualistic: agents can have values without thinking about whether the values are justified, and agents can have values without realizing that they have these values.

Clark and Dudrick do attempt to address this point. They note that “taking a disposition to be justified, supported by reasons” can simply consist of

the feeling of acting in accord with one’s ‘conscience’... The one who takes herself to act in accord with values and not just desires ‘listens’ to her conscience when she judges ‘this is right.’ It is this (largely inchoate) feeling of being justified that is the key to have a value... as opposed to a mere set of desires or dispositions. (Clark and Dudrick 2007: 216)

Here, Clark and Dudrick weaken the initial claim that values are affects that the agent takes to be rationally warranted. Instead, they offer the following interpretation:

Valuing X = having a positive affective orientation toward X and either (1) regarding this affective orientation as justified, or (2) feeling that this affective orientation is in accordance with one’s conscience.

Might this interpretation succeed?



The appeal to conscience does render the Clark/Dudrick view less intellectualistic. However, it threatens to undermine the initial claim that values essentially involve justifications. Nietzsche writes,

The content of our conscience is everything that was regularly demanded of us without reason during our childhood, by persons whom we respected or feared... (WS 52)

Here, Nietzsche explicitly distances conscience from justification: one's conscience is nothing more than another feeling (cf. GM II.24). Nietzsche's dismissive attitude toward conscience is perhaps even clearer in WS 38, where he writes, "The bite of conscience, like the bite of a dog into stone, is a stupidity."

If valuing is having an affect and regarding that affect as in accord with one's conscience, then values are simply affects that are backed by another, higher-order affect. This vitiates the initial claim that values essentially involve a thought about justification. I feel approval toward something, and I feel comfortable or satisfied with that approval. But this satisfaction need have nothing to do with reasons.

Moreover, the Clark/Dudrick view may still seem a bit too reflective. For this interpretation would seem to rule out Nietzsche's claim that we can have values of which we are ignorant. Consider again the philosopher or scientist who, according to Nietzsche, values asceticism. He cannot *approve* of his attraction to asceticism, given that he is ignorant of *having* that attraction. On the Clark/Dudrick view, this would imply that he does not value asceticism.

## 2.4 A new interpretation of Nietzschean values

I have discussed the Richardson and Poellner/Clark/Dudrick views at length because they represent perennially tempting interpretations of Nietzsche on value; indeed, Nietzsche's phrasings positively encourage us to identify values either with drives or with affects. But I hope to have revealed some reasons for resisting these interpretations.

With this groundwork in place, I propose a new interpretation of Nietzsche on value:

An agent values X iff the agent (1) has a drive-induced positive affective orientation toward X, and (2) does not disapprove of this affective orientation.

At this point, we've backed off rather far from the initial Poellner/Clark/Dudrick claim that valuing is having an affective orientation regarded as supported by reasons. Instead, I propose that valuing is having a *drive-induced* affective orientation and *not disapproving* of this orientation.<sup>3</sup> The affective orientation must be *drive-induced* because, as I've explained above, fleeting affective orientations cannot be identified with the agent's values; drive-induced affective orientations, on the other hand, will have the stability and pervasiveness that we associate with values. The proposed interpretation focuses on the *absence of disapproval* rather than the presence of approval for two reasons: it enables the view to account for unreflective values, and it makes the default position valuing the ends of one's drives. I will explain these points in turn.

First, consider again GM III.24's claim that the scientist values asceticism. Certainly, the scientist doesn't *approve* of asceticism in any robust sense—after all, he may never have given a thought to asceticism. But he is unwittingly attracted to asceticism, and he does not disapprove of this attraction. This is enough, on my interpretation, for him to value asceticism.

Second, by focusing on the absence of disapproval rather than presence of approval, we capture the idea that the *default* position is valuing the aims of one's drives. If the view had required approval rather than the absence of disapproval, then there would be an extra step interposed between having a drive toward X and valuing X. Instead, on my proposed account it takes work not to value the aims of one's drives.

I've explained how this new interpretation differs from and improves upon the Poellner/Clark/Dudrick interpretation. Let's now consider how it differs from that of Richardson. First, recall Richardson's claim that values are identified with the ends of our drives. This view seemed problematic: an agent with an affective orientation toward X that he regards as unjustified will count as valuing X. For example, consider the ascetic who consciously maintains that sex is disvaluable, yet has a powerful, drive-induced affective

---

<sup>3</sup> Disapproval can take a variety of forms. It can be manifest in explicit judgments such as "X lacks warrant," "X is not valuable," "my attraction to X is irrational," and so forth. Alternatively, it can consist of the feeling that X is not in accordance with one's conscience. More modestly, it can be a feeling of aversion, doubt, conflict, or dissatisfaction with one's attitude toward X.

orientation towards sexual activity. On Richardson's view, this ascetic values sex, despite his conscious endorsement of the claim that sex is disvaluable. However, according to the account of values offered above, the ascetic will not count as valuing sex. In regarding his positive affective orientation toward sex as unjustified, he will not fulfill the conditions for valuing sex.

In addition, Richardson's view attempted to account for discrepancies between drives and conscious value judgments by introducing a distinction between unreflective values and reflective values: unreflective values are the goals of our drives, whereas reflective values are our consciously espoused principles. I suggested that this strong distinction between two forms of values faces a pair of difficulties: it must explain both why these unreflective values count as *values* (rather than mere dispositions), and how these unreflective values relate to reflective values. Neither of these tasks seemed straightforward.

However, notice that my interpretation avoids these problems. It does not require a sharp distinction between the agent's unreflective values and reflective values. We could use the term "unreflective value" to refer to drive-induced affective orientations that the agent is unaware of having, or to which the agent devotes little thought. And we could use "reflective value" to pick out drive-induced affective orientations that are accompanied by explicit thoughts about approval or justification. However, these forms of valuing are continuous. They are not different kinds of mental event; they are the same kind of mental event, accompanied by different degrees of awareness.

### 3. Drives explain values

At this point, we have an explanation of what values are. Moreover, we have seen how drives and affects relate to values: drives induce affective orientations; when these drive-induced affective orientations are not regarded as unjustified, they count as values.

However, we have not yet examined Nietzsche's claim that drives *explain* the agent's consciously accepted values. Consider a few exemplary passages:

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)

morality is also merely a sign language of the affects... (BGE 187)

Our moral judgments and evaluations...are only images and fantasies based on a physiological process unknown to us. (D 119)

Nietzsche clearly believes that drives explain the agent's consciously espoused judgments about values. Does the proposed view make sense of this?

In a limited way, it does. Schematically, we have a picture of the following sort: a drive induces an affective orientation; so long as this affective orientation is not regarded as unjustified, it counts as a value. The default position is valuing the aims of one's drives.

But surely Nietzsche means something more than this. After all, many consciously espoused values seem to diverge from the aims of our drives. Moreover, Nietzsche suggests that drives can be invoked to explain our conscious thoughts about which values are justified. He writes that a drive

erupts from time to time as reason and passion of mind; it is then surrounded by a resplendent retinue of reasons and tries with all its might to make us forget that fundamentally it is drive, instinct, stupidity, lack of reasons. (GS 1)

So drives somehow lead the agent to perceive certain affective orientations (and actions) as justified. As Nietzsche elsewhere puts it, a drive will "emphasize certain features and lines in what is foreign, in every piece of the 'external world', retouching and falsifying the whole to suit itself" (BGE 230).

Thus, Nietzsche claims that when a drive is active it will not only generate affective orientations toward an action; in addition, it will lead the agent to see this affective orientation as justified. If this is right, then drives will explain values in a very strong sense: by inducing both affects and thoughts about the affects' being justified, drives will *cause* us to value their aims. In the following sections, I will explicate this aspect of Nietzsche's view.

I begin with Schopenhauer, who provides a clear and explicit account of the way in which drives cause us to value their aims. I then explain how drives generate thoughts about justification.

### 3.1 Schopenhauer on the illusions induced by drives

In a chapter entitled “The Metaphysics of Sexual Love,” Schopenhauer offers an intriguing discussion of the way that a particular drive—the sex drive—operates.<sup>4</sup> This account explains how a drive can affect thoughts about justification. Below, I will reconstruct the relevant portions of the account.

Schopenhauer chooses to focus on sexual love because it is among the strongest and most pervasive of human motives:

next to the love of life, it [sexual love] shows itself here as the strongest and most active of all motives, and incessantly lays claim to half the powers and thoughts of the younger portion of mankind. It is the ultimate goal of almost all human effort... (WWR II: 533)

The pervasiveness and force of sexual love should be obvious. But how, exactly, should we account for it?

Schopenhauer’s first claim about love is this: the drive that is responsible for our experience of love does not aim at love, but rather at sexual or reproductive activity.<sup>5</sup> As he puts it, “the true end of the whole love-story, though the parties concerned are unaware of it, is that this particular child may be begotten” (WWR II: 535). For “the ultimate aim of all love-affairs... is nothing less than the *composition of the next generation*” (WWR II: 534). More generally,

---

<sup>4</sup> Katsafanas (forthcoming) contains a brief discussion of this section.

<sup>5</sup> Let me flag a complication at the outset. I argued above that Nietzschean drives incline us to engage in characteristic forms of activity, rather than to realize particular goals or bring about certain achievements. In light of this fact, we should distinguish between a drive’s aiming at *reproduction* and a drive’s aiming at *sexual activity*. Reproduction is a goal or an achievement, rather than a characteristic form of activity. Accordingly, reproduction cannot be the aim of a drive; the drive’s aim is, instead, sexual activity. (Indeed, we might suspect that Schopenhauer is confusing the drive’s *function* with the drive’s aim.) Although Schopenhauer’s treatment of drives elides the distinction between activities and mere goals, his analysis of the way in which the drive impacts reflective thought is unaffected by this oversight. Accordingly, I ignore this point in the text. See also note 7, below.

all amorousness is rooted in the sexual impulse alone, is in fact absolutely only a more closely determined, specialized, and indeed, in the strictest sense, individualized sexual impulse, however ethereally it may deport itself. (WWR II: 533)

In these passages, Schopenhauer claims that love is motivated by a drive that aims at reproduction. We do not have a drive toward love as such; rather, we have a drive toward reproduction, which leads us to experience love.

Of course, we do not experience love as geared solely toward reproduction. Indeed, many individuals who are in love, and who engage in sexual activity, desire not to reproduce. As Schopenhauer puts it, these individuals “abhor... and would to prevent the end, procreation, which alone guides” the drive (WWR II: 540). So Schopenhauer wants to claim both that love is motivated by a drive toward reproduction, and that individuals can be moved by this drive toward reproduction while abhorring reproduction. How are these claims consistent?

Schopenhauer explains that the drive operates by occluding its aim:

the sexual impulse, though in itself a subjective need, knows how to assume very skillfully the mask of objective admiration, and thus to deceive consciousness; for nature requires this stratagem in order to attain her ends. (WWR II: 535)

In other words, the reproductive drive *disguises* its true aim. “However objective and touched with the sublime that admiration may appear to be” (WWR II: 535), what is really aimed at is reproduction.

Indeed, Schopenhauer suggests that the reproductive drive needs to disguise its aim in order to achieve this aim:

Egoism is so deep-rooted a quality of individuality in general that, in order to rouse the activity of an individual being, egoistical ends are the only ones on which we can count with certainty... when the individual is to be active, and even to make sacrifices for the sake of the continuance and constitution of the species, the importance of the matter cannot be made so comprehensible to his intellect...

Therefore in such a case, nature can attain her end only by implanting in the individual a certain *delusion*, and by virtue of this, that which in truth is merely a good thing for the species seems to him to be a good thing for himself, so that he serves the species, whereas he is under the delusion that he is serving himself. In this process a mere chimera, which vanishes immediately afterward, floats before him, and, as motive, takes the place of reality. This *delusion* is *instinct*. (WWR II: 538)

In other words, the reproductive drive presents itself in a guise that will incline the agent to act on it. Precisely because the individual does not have a conscious desire for reproduction (and may have a conscious desire *not* to reproduce), the drive presents its aim (sexual activity resulting in reproduction) in terms that will appeal to the agent. (Schopenhauer claims that this will be an egoistic end, though his argument does not depend on this point.)

In short,

Here, then, as in the case of all instinct, truth assumes the form of a delusion, in order to act on the will. (WWR II: 540)<sup>6</sup>

So, on Schopenhauer's account, the drive has a quite definite end: reproduction. However, in order to achieve this end, the drive generates delusions. It makes us think that we desire to be with a particular person, that this person will offer us complete happiness, that we will attain a total satisfaction from being with this person.

The fact that the reproductive drive's aim diverges from the agent's consciously experienced end is revealed in the *loss* that agents sometimes experience after sex:

[The reproductive drive] casts itself on persons who, apart from the sexual relation, would be hateful, contemptible, and even abhorrent to the lover. But the will of the species is so much more powerful than that of the individual, that the lover shuts his eyes to all the qualities repugnant to him, overlooks everything, misjudges everything, and binds himself forever to the object of his passion. He is so completely

---

<sup>6</sup> Compare the following passage: "Instinct is everywhere an action as if in accordance with the conception of an end or a purpose, and yet entirely without such a conception. Nature implants it, wherever the acting individual would be incapable of understanding the end, or unwilling to pursue it" (WWR II: 540).

infatuated by that delusion, which vanishes as soon as the will of the species is satisfied, and leaves behind a detested partner for life. (WWR II: 555)

More generally,

because the passion rested on a delusion that presented as valuable for the individual what is of value only for the species, the deception is bound to vanish after the end of the species has been attained. ... The individual falls back into his original narrowness and neediness, and sees with surprise that, after so high, heroic, and infinite an effort, nothing has resulted for his pleasure but what is afforded by any sexual satisfaction. Contrary to expectation, he finds himself no happier than before; he notices that he has been the dupe of the will of the species. (WWR II: 557)

Once the drive has attained its aim, the illusions are dispersed, and the agent experiences his beloved as “detested,” wondering why he spent so much effort attaining an end so mundane. That this is not a universal post-sexual experience need not undermine Schopenhauer’s point: he can argue that the reproductive drive sometimes carries individuals along for many a year.

With these points in mind, we can summarize Schopenhauer’s claims as follows:

- (1) The sex drive aims at reproduction.
- (2) However, the individual motivated by the sex drive typically does not desire to reproduce.
- (3) Thus, in order to achieve its aim, the sex drive inclines the agent to experience the pursuit of sex as appealing for reasons unrelated to reproduction. (For example, he sees the lover as affording him boundless happiness or pleasure).
- (4) However, this is an illusion: if the agent were not under the influence of the drive, he would not experience his pursuit of sexual relations as appealing.

Although Schopenhauer focuses on sexual love, he views all drives as operating in this fashion (WWR II: 540).



The key claim is (3). We might wonder how, exactly, a drive is able to affect the agent's experience. Schopenhauer's descriptions imply that the sex drive causes the agent to do at least the following:

- (3a) to overlook disagreeable aspects of potential reproductive partners
- (3b) to find agreeable aspects of potential reproductive partners salient
- (3c) to experience the potential reproductive partner in a different way (as beautiful rather than ordinary, as alluring rather than aversive, etc.)
- (3d) to experience positive affective orientations toward the potential reproductive partner

While Schopenhauer's descriptions provide some illustrations of these points, it may be useful to provide a more detailed and realistic example. To that end, consider Swann's love for Odette in *Remembrance of Things Past*.

Initially, Swann regards Odette as unattractive:

She had struck Swann not, certainly, as being devoid of beauty, but as endowed with a kind of beauty which left him indifferent, which aroused in him no desire, which gave him, indeed, a sort of physical repulsion... Her profile was too sharp, her skin too delicate, her cheekbones were too prominent, her features too tightly drawn, to be attractive to him. Her eyes... seemed to droop beneath their own weight, strained the rest of her face and always made her appear unwell or in a bad mood. (Proust 1982: 213)

Swann soon falls in love with Odette, though, and his perceptions shift: he sees her face as "a skein of beautiful, delicate lines" (244), he remarks on "her brilliant eyes" (254), and he "associat[es] the thought of Odette with his dreams of ideal happiness" (244). In short, she "assumed a new and nobler form" (245). Indeed, Swann sees Odette as the embodiment of the feminine beauty captured in Florentine paintings (244).

Eventually, Swann's love dissipates. He sees again "Odette's pallid complexion, her too thin cheeks, her drawn features, her tired eyes, all the things which...he had ceased to notice since the early days of their intimacy" (415). Thus, he exclaims to himself: "To think

that I've wasted years of my life, that I've longed to die, that I've experienced my greatest love, for a woman who didn't appeal to me, who wasn't even my type!" (415).

At this moment, Odette's features present themselves quite differently: her alluring eyes have become merely tired, her fair complexion pallid. As Schopenhauer would describe it: the sex drive has released its hold on Swann, and the illusions formerly induced by it disappear. Swann now sees Odette in the way that a disinterested observer might see her. He has been "the dupe" of his drive (WWR II: 557).

This process, described in exquisite detail by Proust over the course of some two hundred pages, is a perfect example of what Schopenhauer has in mind. Swann overlooks disagreeable aspects of Odette, finds agreeable aspects salient, experiences Odette in a different way (her eyes appear as alluring rather than droopy, her face is a skein of beautiful lines rather than sharp and unattractive), and experiences a variety of powerful affects toward her. This is what Schopenhauer terms an "illusion."

### **3.2 How drives generate thoughts about justification**

With this in mind, we can see why drives tend to generate thoughts about justification. If I have a positive affective orientation toward X, and if this affective orientation inclines me to experience X in ways that would justify that very affective orientation, then I will be strongly inclined to value X. Consider Swann in the midst of his love for Odette: with the drive active, he sees her as the embodiment of all feminine qualities that he antecedently accepts as justifying love. Seeing her as possessing these qualities, he quite reasonably concludes that his love for Odette is justified. The problem, of course, is that Swann only experiences Odette as possessing these qualities because of the activities of his drive.

I submit that Nietzsche's account of the operations of drives is analogous. Whereas Schopenhauer describes drives as generating "illusions" or "delusions," Nietzsche more often speaks of drives "coloring," "gilding," "lighting," and "staining" the world (see for example GS 7, 139, 152, 301; BGE 186). The terminology is different, but the idea is the

same: drives render certain aspects of experience salient, alter the way that we are inclined to experience the world, and generate affective orientations toward the world.<sup>7</sup>

If this account is correct, then drives *explain* values in a very strong sense. For values are drive-induced affective orientations of which the agent does not disapprove. Drives not only generate the affective orientation, but also incline the agent to see the affective orientation as warranted and therefore worthy of approval. For example, Swann's love for Odette is caused by his reproductive drive, but this reproductive drive also colors his perception and thought in such a way that he experiences his love for Odette as justified by Odette's features. Rather than providing an independent justification for the drive-induced affective orientation, the perception of Odette as beautiful is a product of the drive. It is in this sense that drives surround themselves with a "resplendent retinue of reasons" (GS 1).

#### 4. An objection to the account

Although the reading developed above shows how drives explain reflective judgments about value, it may seem to face a terrible difficulty. Suppose an agent recognizes that his drives are coloring his perceptions and thoughts so that he sees a particular state of affairs as valuable. Wouldn't awareness of these facts render the "illusions" or "colorings" generated by the drive unconvincing? Put differently, how could an agent both recognize that he judges X valuable merely because his drives are coloring his perception of X, and continue to judge X as valuable?

Sebastian Gardner presents this question in a particularly pressing form:

---

<sup>7</sup> For further arguments to this effect, see Katsafanas (forthcoming). It is worth noting that there is an important difference between Schopenhauer and Nietzsche on drives, which shows up in the terminology that they use to describe the effects of drives. Consider the sex drive. Schopenhauer assumes that the sex drive's true end is reproduction, whereas sexual attraction and love are mere illusions induced by the drive so as to incline the agent toward reproduction. But we might pose several questions at this point: what makes reproduction, rather than sexual attraction or love, the true aim of the drive? Why think that sexual attraction and love merit the label "illusions", whereas reproduction is the reality? More generally, given the pervasive impact of drives on affect and conscious thought, why think that this illusion/reality distinction is tenable? Though I lack the space to address this point in detail, it is important to note that Nietzsche speaks not of illusions but—as I note above—of colorings, stainings, and so forth. Nietzsche is not committed to the idea that the drive's aim is somehow uniquely real; indeed, the affective orientations induced by the drive are, for certain explanatory purposes, more relevant and more important than the drive's aim. See, for example, GS 44, where Nietzsche writes, "*Supposed motives*.—Important as it may be to know the motives that actually prompted human conduct so far, it may be even more essential to know the fictitious and fanciful motives to which men ascribed their conduct. For their inner happiness and misery has come to men depending on their faith in this or that motive—not by virtue of the actual motives. The latter are of second-order interest" (GS 44).

the Nietzschean subject lacks any rational warrant for regarding his valuation as anything more than the expression of a natural force; and this notion—that one could take one’s values, not just when viewing oneself sideways on but also *in the very act of* legislating and endorsing them, to be nothing more than the causal effects of pre-normative psychological forces—encounters a problem of sheer inconceivability for subjects in whom the taste for justification is well established. (Gardner 2009: 16)

In other words, if I believe that my values are nothing more than the expression of “pre-normative psychological forces”—drives—then it is inconceivable that I could persist in taking myself to have any rational warrant for maintaining these values. I cannot both accept that I value X *merely* because I have a drive toward X, and maintain that my valuation of X is justified: the thought that the valuation is merely drive-induced undercuts the idea that it might be justified.

If Gardner is correct, then Nietzsche’s account of value is doomed. Below, I will argue that Nietzsche in fact has a pair of responses to this form of objection.

#### 4.1 An initial response

In essence, Gardner claims that once an agent recognizes that he values X merely because of the operations of a drive toward X, the agent will cease to see his valuation of X as justified. Now, the first thing to notice about Gardner’s objection is that it seems rather overstated: although I can reflectively judge that I am only experiencing a valuation as warranted because of the peculiarities of my psychology, this judgment need not lead me to stop experiencing the valuation as warranted. To illustrate this point, let’s consider two different cases: one in which I explicitly take the psychological basis of my valuation to be unproblematic, and another in which I take the psychological basis to be objectionable.

Consider again HH I.226:

The fettered spirit takes up his position, not for reasons, but out of habit; he is a Christian, for example, not because he has knowledge of the various religions and has chosen between them... he encountered Christianity... and adopted it without

reasons, as a man born in a wine-producing country becomes a wine drinker. Ask a fettered spirit for his reasons against bigamy, for example, and you will learn whether his holy zeal for monogamy is based on reasons or on habit. (HH I.226)

Focus, first, on the wine-drinking example. I can recognize that I only have a taste for wine because I have been exposed to it repeatedly, and have—under the pressure of judgments about its value, out of an attempt to appreciate that which others appreciate—acquired tastes for it. I can recognize that my valuation of wine is wholly contingent, based merely on this affective orientation. Nonetheless, given that my affective orientation toward wine is what it is, I take my valuation of it to be justified. In this case, I can simultaneously recognize the contingent, pre-rational basis of this evaluation and persist in my positive evaluation of wine.

Contrast this with Christianity. Suppose I value Christian practices, but come to believe that I accept these practices only because my upbringing has instilled in me various dispositions, sensibilities, and so forth. This *does* seem to undermine the valuation; this valuation seems to demand a more robust, less contingent form of valuation. It requires a backing that is more than merely psychological.

These examples indicate that Gardner's objection must be modified. There is no *general* problem with taking a valuation to be drive induced while simultaneously regarding it as justified. The problem arises only for specific valuations: it arises only when we judge that the contingency of a valuation is in itself objectionable.

Consequently, we could respond to Gardner's objection by claiming that Nietzsche wants us to treat *all* of our values like the valuation of wine: rather than thinking that our valuations demand a non-psychological backing, we can come to regard drive-based, psychological backing as sufficient.<sup>8</sup> While I think this line of response has considerable merit, I will not pursue it here. For there is another response: we can answer Gardner's objection by appealing to the recalcitrance of drive-induced "illusions."

---

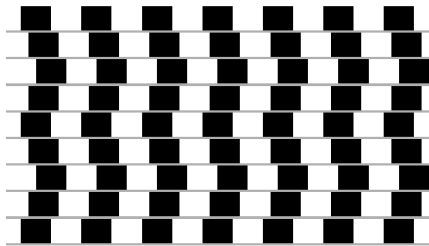
<sup>8</sup> Simon de Beauvoir offers a clear statement of what this form of valuing would involve: "In the earthly domain all glorification of the earth is true as soon as it is realized. Let men attach value to words, forms, colors, mathematical theorems, physical laws, and athletic prowess; let them accord value to one another in love and friendship, and the objects, the events, and the men immediately *have* this value; they have it absolutely" (Beauvoir 1976: 157-8). She continues, "And in fact, any man who has known real loves, real revolts, real desires, and real will knows quite well that he has no need of any outside guarantee to be sure of his goals; their certitude comes from his own drive" (159).

## 4.2 The recalcitrance of illusion

Consider again Gardner’s objection: I cannot both see my valuation of X as merely drive-induced and continue to regard my valuation of X as justified. This would be, Gardner claims, a “sheer inconceivability.”

I will argue that Schopenhauer’s account of drives gives us a way of answering Gardner’s objection. Schopenhauer explicitly states that I can recognize that my love lacks rational warrant—I can recognize that my love is merely induced by a drive whose end that I deplore—without this disrupting the love. Recognition of the valuation’s physiological basis doesn’t dispel the perception of my love as warranted.

Why is this? Schopenhauer quotes with approval an old saying: “what is endowed with neither reason or moderation cannot possibly be ruled by reason” (WWR II: 557). The drive is, Schopenhauer maintains, a pre-rational structuring of thought and affect. This structuring of thought and affect is not subject to direct reflective control; the agent cannot simply decide to eliminate it. As a result, the illusions induced by a drive are as resistant to dispersal as a typical perceptual illusion. Consider a simple example: the café-wall illusion, depicted below.



Each of these black and white boxes is identical in size, and consequently the gray horizontal lines are perfectly straight. Nonetheless, we perceive the boxes as unequal in size and the horizontal lines as bent. One can verify the straightness of the horizontal lines and the equality of the blocks with a ruler; nonetheless, even after doing so one will continue to experience them as bent and unequal.

No amount of reflective judgment will eliminate this illusion. Just so, Schopenhauer and Nietzsche claim, with the “illusions” induced by drives. The processes listed as (3a)-(3d) are resistant to reflective amelioration: even if I recognize that they are occurring, and regard

their occurrence as unwarranted or otherwise problematic, I typically won't be able to eliminate them.

The fact that drive-induced "illusions" are resistant to reflective amelioration enables us to answer Gardner's objection. The point is this: I can simultaneously see that I experience my valuation of X as warranted only because of the distorting effects of my drives, and remain incapable of shedding the perception of my valuation of X as warranted.

To clarify this point, let's distinguish:

- (1) The agent's drive-induced affective orientation toward X, which strongly inclines the agent to see judgments such as "X is valuable" as warranted, in part by altering the way in which he perceives the facts that ground his judgments about value.
- (2) The agent's reflective judgments of the form "X is valuable."

(1) and (2) can come apart: it is conceivable that an agent could have a powerful affective orientation toward X, to be strongly inclined to see judgments such as "X is valuable" as warranted, and nonetheless continue to maintain—in cool moments—that X is not valuable. Swann again provides an excellent example: even while his love for Odette is strong,

He was well aware that his love did not correspond to anything outside itself...he realized that Odette's qualities were not such as to justify his setting so high a value on the hours he spent in her company. And often, when the cold government of reason stood unchallenged in his mind, he would readily have ceased to sacrifice so many of his intellectual and social interests to this imaginary pleasure. (Proust 1982: 258)

When "the cold government of reason" is dominant, Swann can tell himself that Odette isn't worthy of love. He can recognize that although he experiences her as worthy of love, this valuation is not justified: it does not "correspond to anything outside itself." Although he understands this perfectly, his perceptions of Odette remain unchanged. So, although Swann has (1), he lacks (2): he reflectively judges that Odette is not loveable. Nonetheless, he continues to experience powerful affective orientations toward Odette, which strongly dispose him to put such thoughts out of his mind, and convince himself that she really is

loveable.

With these remarks on affective orientations at hand, we can consider the way in which drives strongly dispose agents to make reflective judgments about value. The idea is fairly simple: if I have a strong affective orientation toward X, which inclines me to see X as attractive and hence as worthy of pursuit, then I will be inclined to see valuation of X as warranted. My perceptual experiences will support this valuation: when I do reflect on the justification for X, I will tend to see justifications for that valuation everywhere. Swann reflects on whether his valuation of Odette is justified; his drive inclines him to see Odette as alluring, attractive, capable of providing him with complete happiness. These are the sorts of attributes that would justify his love; hence, he persists in his valuation.

Of course, Swann can step back, and see that he experiences Odette as possessing these positive attributes solely as a result of the sex drive's activity. He can recognize that, were his drive to find some other object, he would cease to find Odette attractive and alluring. But this isn't something over which he has direct control; he cannot simply turn these perceptual orientations on and off. Thus, he is in the curious position of realizing that his valuation of Odette is merely drive-induced—that he experiences Odette as loveable merely because of the way that his drive modifies his perceptual experiences and affects—and nevertheless these recognitions have no tendency to alter the character of his experiences. Even while recognizing that his perceptions of warrant are drive-induced, he continues to see them as perceptions of warrant.

This is why Gardner's objection is unsuccessful. Far from it being a "sheer inconceivability" for an agent simultaneously to view her valuation as induced by a drive and justified, we do this all the time. Nothing is more ordinary than seeing our valuations as induced by our own affects, being incapable of eliminating the affects in question, and hence persisting in the valuations.

## **5. Conclusion: the continuity between drives, affects, and values**

We began by asking how Nietzschean drives and values relate to one another. We can now offer the following account:



- (1) An agent values X iff the agent (i) has a drive-induced positive affective orientation toward X, and (ii) does not disapprove of this affective orientation.
- (2) If an agent has a drive-induced affective orientations toward X, then the agent's perceptions of X will typically be structured in a way that strongly inclines her to see judgments of the form "X is valuable" as warranted.
- (3) These drive-induced structurings of perception are resistant to reflective amelioration.

This account explains Nietzsche's two key remarks about value: that drives include valuations and that drives explain reflective judgments about value. As (1) makes clear, the default position is for agents to value the aims of their drives; valuations and drives will diverge only when the agent disapproves of the drive's aim. Second, claims (2) and (3) show the sense in which drives explain reflective judgments about value; individuals typically experience the world in ways that strongly incline them to judge that their drives' aims are valuable.

I will conclude by considering how this interpretation relates to the other accounts reviewed above. My view incorporates aspects of the Richardson and Poellner/Clark/Dudrick interpretations. Richardson is correct that valuing is rooted in drives; but the drives constitute valuing precisely because they generate affective orientations, which incline the agent to see pursuit of the drive's aim as warranted. Poellner, Clark, and Dudrick are correct that valuing is rooted in affects; but the relevant affects are the aforementioned drive-induced affective orientations, which differ significantly from discrete, humdrum emotions.

My proposed interpretation also provides an account of the connection between drives and affects, on the one hand, and reflective judgments about value, on the other. Return, once again, to Richardson's claim that Nietzsche distinguishes reflective values (our consciously espoused principles) and unreflective values (the ends of our drives). We can weaken this claim: Nietzsche isn't *distinguishing* these two forms of valuing so much as showing that they are *continuous*. There is continuity between having a drive toward X and reflectively espousing a conscious principle such as "X is good." In particular, the drive toward X will generate affective orientations toward X, and these affective orientations will, in turn, strongly incline us to regard principles such as "X is good" as warranted. These

pressures will be resistant to reflective amelioration. Consequently, having a drive toward X and valuing X will tend to go hand in hand: only in aberrant cases will they come apart.

But the continuity between reflective and unreflective values is still stronger: reflective thoughts about value and affective orientations induced by drives occupy analogous functional roles. Consider how we understand reflective values: values are principles that structure our deliberations, inclining us to judge that pursuit of the valued object is justified. Suppose we accept something like the following account of value:

To value X is, in part, to be disposed to make conscious judgments of the form “X is valuable,” to experience pursuit of X as warranted, to have a positive affective orientation toward X, to pursue X, to admire other agent’s pursuit of X, and so on.

Drive-induced affective orientations have an analogous functional role. They dispose us to make conscious judgments of the form “X is valuable,” to experience pursuit of X as valuable, to have a positive affective orientation toward X, to admire other agents’ pursuits of X, and so on. Thus, the “illusions” or drive-induced affective orientations structure deliberation just as effectively as do explicit, consciously entertained principles.

Significantly, these affective orientations structure our orientations *regardless of whether we regard them as justified*. The perceptual illusion described above demonstrates that structurings of perception are sometimes immune to reflective amelioration. As the passages from Proust illustrate, affective structurings are analogous. This is why it makes sense to view these drive-induced affective orientations as *values*: they occupy functional roles that are analogous to those of reflectively espoused evaluative principles.

In sum, although Nietzsche’s assertion that drives *include* and *explain* values is initially puzzling, the view turns out to make good sense. Drives include values in the sense that values just are drive-induced affective orientations of which the agent does not disapprove. Drives explain values in the sense that drive-induced affective orientations strongly incline the agent to regard judgments of the form “X is valuable” as justified.

## Translations Used

When quoting from Nietzsche's works, I use the translations by Walter Kaufmann and R.G. Hollingdale. In some cases I have modified the translation.

## Works Cited

- Beauvoir, S. de (1976), *The Ethics of Ambiguity*, trans. B. Frechtman (New York: Kensington Press).
- Clark, M. and Dudrick, D. (2009), "Nietzsche on the Will: An Analysis of BGE 19," in Gemes and May (2009).
- Gardner, S. (2009), "Nietzsche, the Self, and the Disunity of Philosophical Reason," in Gemes and May (2009).
- Gemes, K. and May, S. (2009), *Nietzsche on Freedom and Autonomy* (Oxford: Oxford University Press).
- Janaway, C. (2006), "Nietzsche on Free Will, Autonomy and the Sovereign Individual," *Aristotelian Society Supplementary Volume* 80: 339-357.
- (2007), *Beyond Selflessness: Reading Nietzsche's Genealogy* (New York: Oxford University Press).
- Katsafanas, P. (forthcoming), "Nietzsche's Philosophical Psychology," in Richardson and Gemes (forthcoming).
- Leiter, B. and Sinhababu, N. (2007), *Nietzsche and Morality* (New York: Oxford University Press).
- Poellner, P. (2007), "Affect, Value, and Objectivity," in Leiter and Sinhababu (2007).
- Proust, M. (1982), *Remembrance of Things Past*, Volume I, trans. C.K. Scott Moncrieff and T. Kilmartin (New York: Vintage).
- Richardson, J. (2004), *Nietzsche's New Darwinism* (New York: Oxford University Press).
- (forthcoming), "Nietzsche on Life's Ends," in Richardson and Gemes (forthcoming)
- Richardson, J. and Gemes, K. (forthcoming), *The Oxford Handbook on Nietzsche* (Oxford: Oxford University Press).
- Schopenhauer, A. (1969), *The World as Will and Representation*, trans. E.F.J. Payne (New York: Dover).