https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

MORALITIES ARE A SIGN-LANGUAGE OF THE AFFECTS*

By Brian Leiter

This essay offers an interpretation and partial defense of Nietzsche's idea that moralities and moral judgments are "sign-languages" or "symptoms" of our affects, that is, of our emotions or feelings. According to Nietzsche, as I reconstruct his view, moral judgments result from the interaction of two kinds of affective responses: first, a "basic affect" of inclination towards or aversion from certain acts, and then a further affective response (the "meta-affect") to that basic affect (that is, sometimes we can be either inclined towards or averted from our basic affects). I argue that Nietzsche views basic affects as *noncognitive,* that is, as identifiable solely by how they feel to the subject who experiences the affect. By contrast, I suggest that meta-affects (I focus on guilt and shame) sometimes incorporate a *cognitive* component like belief. After showing how this account of moral judgment comports with a reading of Nietzsche's moral philosophy that I have offered previously,[1] I conclude by adducing philosophical and empirical psychological reasons for thinking that Nietzsche's account of moral judgment is correct.

I. Sign-Languages and Symptoms

My title comes from section 187 of Nietzsche's *Beyond Good and Evil,* though the theme in question—that moralities are "a sign-language of

* I have benefited from discussion of the issues treated in this paper with the participants in the Spring 2011 seminar on "Nietzsche and Moral Psychology" at the University of Chicago: Nir Ben-Moshe, Jaime Edwards, Roger Eichorn, Guy Elgat, Michael Forster, Simon Gurofsky, Peter Kail, Alex Langlinais, and David Showalter. Teaching Nietzsche's *Genealogy* with Peter Kail at Oxford in autumn 2011 was also very helpful in clarifying my thoughts on several issues treated here. An earlier version of this paper was presented as the Bernd Magnus Lecture to the Department of Philosophy at the University of California at Riverside on March 7, 2012. I am grateful to the audience there for helpful questions and discussion; I should mention in particular David Glidden, Pierre Keller, Samantha Matherne, Jozef Müller, and Howard Wettstein. The current version benefited from critical discussion with João Constâncio and his colleagues and students at the Nietzsche International Lab at the New University of Lisbon; from discussion with Justin Coates and with the other contributors to this volume; from the comments of an anonymous referee for this journal and from the editor, David Schmidtz; and from written commentary by Maudemarie Clark and David Dudrick, and discussion with them and the audience at an invited session on "Nietzsche, Moral Psychology and Empirical Psychology" at the Central Division meeting of the American Philosophical Association in February 2013.

[1] Brian Leiter, *Nietzsche on Morality* (London: Routledge, 2002).

doi:10.1017/S0265052513000113
© 2013 Social Philosophy & Policy Foundation. Printed in the USA.

the affects"—is a pervasive one in Nietzsche's work throughout the 1880s and central to his moral psychology. In *Daybreak,* he declares that "our moral judgments and evaluations . . . are only images and fantasies based on a physiological process unknown to us."[2,3] In *The Gay Science,* he suggests that, "[a]nswers to the questions about the *value* of existence . . . may always be considered first of all as the symptoms of certain bodies."[4] Again, in *Beyond Good and Evil,* he tells us that a philosopher's "morality" simply bears "decisive witness to . . . the innermost drives of his nature."[5] In the *Genealogy,* Nietzsche famously claims that "our . . . values . . . grow from us with the same inevitability as fruits borne on the tree."[6] In *Twilight of the Idols,* Nietzsche remarks that "judgments of value . . . have value only as symptoms."[7] And in a *Nachlass* note of 1885–86, he says: "Moral judgments [are] symptoms and sign languages which betray the process of physiological prosperity or failure."[8]

Nietzsche has two main metaphors for describing moralities or systems of value—that of *sign-language* (*Zeichensprache*) and of *symptom* [*Symptome*]—and two main idioms in which to explain the referent of the sign-language or cause of the symptom, one psychological involving affects (*Affekten*), feelings (*Gefühle*), or drives (*Triebe*), and the other physiological. The physiological idiom is undoubtedly important to Nietzsche, influenced as he was by the German Materialist movement in Germany in the mid-nineteenth century,[9] but it seems equally clear that, apart from calling attention to the possibility of physiological explanations for evaluative orientations, Nietzsche makes no real intellectual contribution to this kind of explanation.[10] Like Freud, Nietzsche plainly thinks that some class of mental phenomena are physically explicable, though, unlike Freud, he was fairly explicit in rejecting any type-identity of mental and physical states of the person.[11] But Nietzsche is

[2] I have drawn on English translations of Nietzsche's works by Walter Kaufmann, R. J. Hollingdale, or Maudemarie Clark and Alan Swensen and then made modifications based on Friedrich Nietzsche, *Sämtliche Werke: Kritische Studienausgabe in 15 Bänden,* ed. G. Colli & M. Montinari (Berlin: de Gruyter, 1980); where there is no existing English edition, the translation is my own. Nietzsche's works are cited as follows, unless otherwise noted: roman numerals refer to major parts or chapters in Nietzsche's works; arabic numerals refer to sections, not pages. I use the standard Anglophone abbreviations for Nietzsche's works, as follows: *The Antichrist* (*A*); *Beyond Good and Evil* (*BGE*); *The Birth of Tragedy* (*BT*); *Daybreak* (*D*); *Ecce Homo* (*EH*); *The Gay Science* (*GS*); *On the Geneology of Morals* (*GM*); *Thus Spoke Zarathustra* (*Z*); *Twilight of the Idols* (*TI*); *The Will to Power* (*WP*).

[3] Nietzsche, *D,* 119.

[4] Nietzsche, *GS,* "Preface," 2.

[5] Nietzsche, *BGE,* 6.

[6] Nietzsche, *GM,* "Preface," 2.

[7] Nietzsche, *TI,* "Problem," 2.

[8] Nietzsche, *WP,* 258.

[9] See Leiter, Nietzsche on Morality, 63–71.

[10] He acknowledges as much in the "Note" at the end of *GM* I when he calls for a prize to encourage physiologists and doctors to study the effects of different values on persons.

[11] Leiter (2002: 24–25).

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

like Freud in another respect, namely, that the explanatory idiom in which he does most of his work is a *psychological* one, and it is on that idiom we shall focus here.

What could it mean to diagnose particular evaluative judgments, or whole systems of value, as a "sign-language" or "symptom"? The idea of a "symptom" presents the more straightforward case. Consider: my *sore throat* is the symptom of a *viral infection.* That means: the symptom (the soreness of the throat) is *caused by* the virus. So if moral judgments are *symptoms* of certain psychological states, that means those states are *causes* of the moral judgments. Of course, the relationship is not *simply* causal: rather the symptoms count as *evidence* for their cause, because they reveal something about the psychological states that give rise to them.

Sign-languages are, at first blush, a bit different. A sign-language is some system of symbols or signs that have semantic content, that is, have some *meaning* in virtue of *representing* something else.[12] If I raise just my pointer and middle finger, that *means* either peace or victory. The referent of a sign with representational content need not, however, be the cause of that content, though there is certainly one familiar, contemporary view of semantic content in which that would be the case. But like a symptom, the sign does stand in some kind of meaningful inferential relation with its referent (or cause): in both cases, the referent, or cause, is expressed by the symptom or sign. Since Nietzsche uses the metaphors of "symptom" and "sign-language" interchangeably, and since there is no evidence he had a well-developed semantic view, it is natural to understand them the same way. Thus, in what follows, I will assume that the claim that a morality is a "symptom" or "sign-language" of X means that X is the *cause* of the morality (or the cause of some person making a particular evaluative or moral judgment) and, moreover, it is a cause whose existence can be correctly inferred from the symptom or sign.[13]

What, then, of the psychological causes on offer? Nietzsche refers, as noted above, to just three: affects, feelings, and drives. I take *affects* and *feelings* to refer usually to the same kind of mental state for Nietzsche,[14] while I will construe drives, following Paul Katsafanas,[15] as *dispositions to have affective responses under certain conditions.* We will consider each of

[12] There are nonrepresentationalist views of semantic content (think Robert Brandom or Huw Price, for example), but there is no reason, of course, to think Nietzsche had such a view in mind, or, indeed, that he had loyalty to any particular theory of meaning. It is thus best to interpret his remarks in the most natural, "ordinary-language" way that still allows us to make sense of his terminology.

[13] I will thus assume that to the extent a symptom or sign *expresses* a meaning, it does so in virtue of how it is caused by the underlying item of which it is a sign or symptom.

[14] There are occasional exceptions, e.g., Nietzsche, *BGE* 19, but these are somewhat anomalous.

[15] Paul Katsafanas, "Nietzsche's Philosophical Psychology," in *A Handbook of Nietzsche,* ed. Ken Gemes and John Richardson (Oxford: Oxford University Press, forthcoming 2013).

these in turn, before taking up the reasons for thinking that, suitably construed, moralities *really are* symptomatic of our affects.[16]

II. Feelings and Affects

In *Daybreak*, "moral feelings" (*Moralische Gefühle*) are said to be inculcated when "children observe in adults inclinations for and aversions to certain actions and, as born apes, *imitate* these inclinations [*Neighungen*] and aversions [*Abneighungen*]; in later life they find themselves full of these acquired and well-exercised affects [*Affekten*] and consider it only decent to try to account for and justify them."[17] We may bracket for the moment the astute concluding observation—about the impulse to supply post hoc rationalizations for evaluative judgments produced by a nonrational mechanism—in order to focus for now on what this passage tells us about Nietzsche's conception of affects and moral judgment.

Moral feelings or affects, in the passage under consideration, are equated with "inclinations for and aversions to certain actions," or, more precisely, with the mental state, whatever it is, that motivates one to perform certain actions or avoid certain other actions.[18] Nietzsche's ontology of the mind thus includes (unsurprisingly) mental states that are *motivationally effective,* at least in the sense that they produce what I will call henceforth *motivational oomph* (or *push*); that is, they incline toward or avert away from certain acts, even if they are not ultimately successful in producing action. In this respect, Nietzsche's usage is quite intelligible to us: we typically presume that *affects* or *feelings* are characterized, in part, by their ability to produce *motivational oomph,* and, indeed, the fact that they do so is one of the main points thought to count in favor of metaethical views that understand moral judgments to involve essentially the expression of feelings (a point to which we will return).

It is equally a part of ordinary usage to take feelings to have another crucial characteristic: namely, that they are phenomenologically distinctive. There is something *it feels like* to be inclined to stop the child from sticking his hand in the fire, and there is something *it feels like* to be inclined to avoid killing a child. A *noncognitivist* view of affects claims that

[16] Of course, many different things might be symptomatic of our affects: e.g., our facial expressions, our behaviors. Nietzsche thinks, however, that moralities reveal things about our affects and drives that are especially important and which we would not otherwise recognize *unless* we interpreted moral judgments as expressive of affects. This will become clearer in the discussion that follows, below.

[17] Nietzsche, *D,* 34.

[18] There is a certain affinity here to aspects of Stoic moral psychology, with which Nietzsche, as a classicist, was certainly familiar, though he rejects the ultimately rationalist elements of the Stoic view. Tad Brennan, "Stoic Moral Psychology," in *The Cambridge Companion to the Stoics,* ed., B. Inwood (Cambridge: Cambridge University Press, 2003) is a useful review.

The relationship between Nietzsche's view and that of the Stoics requires independent treatment.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

they can be fully individuated by their distinctive phenomenal feel; a cognitivist view denies that, claiming, instead, that to individuate the affect one also needs to consider some aspect of its *cognitive* (i.e., truth-evaluable) content, such as a belief. Cognitivist views of emotions, notoriously, are forced to deny that human infants can have emotions, since they lack the concepts necessary for having truth-evaluable judgments.[19] That infants cannot have emotions might seem a *reductio* of the cognitivist position, at least to anyone who has ever cared for an infant. But less anecdotally, cognitivism about the emotions runs up against a large body of psychological research showing that, as Jesse Prinz puts it, "emotions can arise without judgment, thoughts, or other cognitive mediators."[20] While it may be doubtful that qualitative feel can distinguish all emotional states (Prinz's example is differentiating "anger" from "indignation",[21] that is not necessarily pertinent to Nietzsche's concerns. For Nietzsche locates the *affective* source of moral judgment, in the first instance, in fairly basic or simple mental states of *inclination* and *aversion,* and it seems quite plausible that there is something it feels like to be *inclined toward* X or *averted away from* Y, even if the relevant *qualia* might be thought inadequate to pick out all conceptual nuances we might want to apply to such cases. It is also possible, of course, that our conceptually nuanced distinctions between, say, *anger* and *indignation* are explanatorily otiose with respect to the *motivational oomph* associated with the feeling in question. (The crucial moral emotion of "guilt" presents a special case, however, to which we will return shortly.)

Perhaps more pertinent is the other problem that is supposed to afflict noncognitivist views of feelings or affects, namely, how to account for our propensity to think emotional responses can be assessed as warranted or not.[22] Even in infants, of course, emotional responses have causal triggers: the infant *sees* a scary monster and reacts by crying in fear. The infant's *fear* has an *intentional content:* it is *about* something, namely, the scary monster. We might believe the infant is wrong to be afraid (i.e., what the infant sees is not dangerous, since it is only make-believe), but that does not change the fact that the infant experiences upset and fear. I believe that a child is about to stick his hand in the flame, and so I *feel inclined* to stop him. "Feel inclined" understates the character of the feeling, of course: *I feel I must* stop him. But the fact that there is a causal relationship between a mental state with cognitive content and an affective response, does not show that to *identify* the affect one must understand it to be constituted, in whole or even in part, by a *cognitive* component. We need to account for the *intentionality* of emotions, to be sure, but that

[19] Cf. John Deigh, "Cognitivism in the Theory of the Emotions," *Ethics* 104 (1994): 824–54.
[20] Jesse Prinz, *The Emotional Construction of Morals* (Oxford: Oxford University Press, 2007), 57.
[21] Ibid., 52.
[22] *Cf.* Prinz, ibid., 60ff.

does not mean we have to identify emotions with *judgments* or, more generally, with propositional attitudes like belief and desire. It might suffice to introduce a kind of explanatorily primitive mental state, as Peter Goldie famously proposes, like "feeling toward," one that capture both the *intentionality* of emotional states *and* their distinctive qualitative character.[23]

So feeling inclined to do X or averse to doing Y can have a distinctive phenomenal character, even if the causal trigger for the feeling is a false belief of the agent with the distinctive feeling. When we criticize the resulting emotion, perhaps, then, we are making a different normative judgment, something like: a correct belief about the facts ought not cause such an affective response. Consider: the infant falsely perceives (without any conceptual content) the family dog as dangerous, and so bursts into tears. We think the infant is *mistaken,* because we know that the family dog is not dangerous. But we do not criticize the infant for having an unwarranted affective response, rather we console the infant. Why? Perhaps because we do not think the infant is *epistemically responsible* for its false cognition of the family dog as dangerous. But when I react with anger to a perceived slight by the store clerk that, in reality, was nothing of the kind, I can be criticized not because anger is anything other than noncognitive, and not because my "belief that I have been slighted" is a causal trigger for my anger, but because I am epistemically culpable for having a false belief, one that produces unhappy consequences (my anger, a fight with the clerk at the store, an unpleasant scene, and so forth).

Yet it seems I can also be criticized for responding with anger even when my belief was correct, i.e., when the store clerk *really was* rude. How can that criticism make sense if the cognitive belief is just a causal trigger for an evaluative/affective response that has no cognitive content? The obvious answer is that such judgments are, themselves, the expression of other noncognitive attitudes or feelings, such as an inclination towards nonaggressive responses to insults or an aversion to aggressive ones (we might call the latter the "Christian attitude"). Of course, if we view such criticism as making a cognitive, rather than noncognitive, claim, then we will be forced towards an error-theoretic interpretation of such judgments (that is, we will have to interpret all such judgments as false). But on the general Nietzschean view, it is surely natural to treat such normative criticisms as, themselves, yet more expressions of affects or feelings.[24]

I belabor these points because I want to argue that Nietzsche's view of affects or feelings is basically *noncognitivist:* he thinks our basic affects of inclination and aversion are marked by a distinctive conscious, qualita-

[23] Peter Goldie, *The Emotions: A Philosophical Exploration* (Oxford: Oxford University Press, 2002), 19.

[24] As I discuss below, I actually think we should resist ascribing any *semantic* view to Nietzsche, so the claims in this paragraph should, strictly speaking, be treated as *explanatory* of moral judgment, rather than as part of a semantics of moral judgment.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

tive feel. And he thinks such qualitatively distinctive *feels* are the causal root of our moral judgments, even if those fully articulated judgments are influenced by other factors, to which we will turn momentarily. Given that Nietzsche focuses mainly on *inclination* and *aversion toward* as the crucial affective states, the claim that they can be individuated in noncognitive terms is especially plausible. But we can now also see how the social practice of assessing such feelings as warranted or unwarranted can be compatible with the noncognitive character of the feeling: to deem them warranted or unwarranted is either (1) to render an epistemic judgment about the cognitive judgment that is the causal trigger for the noncognitive state, or (2) to express a meta-affect (that is, an affect *about* an affect) about someone else's affective response to a causal trigger.

III. Drives

The other key element of Nietzsche's ontology of the mind, for purposes of explaining motivation and action, is the notion of a *drive,* which, following Katsafanas, we will treat as a *disposition to have a particular affective response under certain circumstances.*[25] Thus, the sex drive would be a disposition to become sexually aroused in the presence of an attractive member of the opposite or same sex.[26] The important relationship between drives and affects is usefully illustrated by this passage from *Daybreak:*

> The same drive evolves into the painful feeling of *cowardice* under the impress of the reproach custom has imposed upon this drive; or into the pleasant feeling of *humility* if it happens that a custom such as the Christian has taken it to its heart and called it *good.* That is to say, it is attended by either a good or a bad conscience! In it itself it has, *like every drive,* neither this moral character nor any moral character at all, nor even a definite attendant sensation of pleasure or displeasure: it acquires all this, as its second nature, only when it enters into relations with drives already baptized good or evil or is noted as a quality of beings the people has already evaluated and determined in a moral sense. . . .[27]

In the particular example on offer from Nietzsche, the same *drive* is said to have the potential to give rise to two different moral feelings, that of *cowardice* (which has an unpleasant valence) or *humility* (a pleasant valence), depending on the cultural context. The drive in question must be something like, *a disposition to avoid offending dangerous enemies,* which, if expe-

[25] Katsafanas, "Nietzsche's Philosophical Psychology."

[26] The discharge of that drive would then involve acting on the sexual arousal or urge, but the drive can manifest itself in an affective response, even when it is not satisfied.

[27] Nietzsche, *D,* 38.

rienced by a Homeric Greek would then give rise to feelings of self-contempt for being a coward and, if experienced by a Christian, would be experienced as the admirable virtue of humility. Cultures, partly through the mechanisms of parental inculcation already noted (as well as concurrent social pressures), teach individuals to have particular affective responses to the very same drive. Notice, however, that we now have two layers of affects here: first, there is the affect of *aversion towards offending dangerous enemies* which is produced by the drive itself, but *then* there is the distinctively *moral affect* of feeling ashamed (as the Homeric Greek does) or proud (as the Christian does) of that affective response. The *moral affect*, then, is more complicated than what I will henceforth call "the basic affect" of inclination or aversion. The *feeling of being a coward*, on this account, represents the combination of a *feeling of aversion towards offending a dangerous enemy*, conjoined with a meta-feeling of contempt or disgust for having that original feeling. The meta-feeling is, at bottom, a feeling of aversion away from the basic affect, though perhaps to individuate it correctly we will need to add some kind of *belief* about why that basic feeling of aversion is contemptible: e.g., the belief that Homeric men slay their offending enemies, rather than cower before them. Perhaps this is also what Nietzsche is getting at when he writes: "[B]ehind feelings there stand judgments [*Urtheile*] and evaluations [*Werthschätzungen*] which we inherit in the form of feelings (inclinations, aversions). The inspiration born of a feeling is the grandchild of a judgment—and often of a false judgment!"[28] So, for example, the Christian judges that a good person *chooses* to display humility, a trait admirable in the eyes of God, and thus when the Christian experiences the basic affect of aversion towards offending a dangerous enemy, he then experiences the meta-affect of a positive inclination or valence towards that basic affect: he is proud of his humility. The Christian then inculcates those same affective responses in his children, so that their feelings of inclination and aversion are traceable back to a false judgment, that is, the judgment that there is a God who thinks humility is a virtuous trait, as well as the false belief that the "humility" of the Christian is a free choice, rather than a reaction he cannot avoid having.[29] As Nietzsche observes, "we are all irrational" in that we "still draw the conclusions of judgments we consider false, of teachings in which we no longer believe—our feelings make us do it,"[30] a claim well-supported by the recent empirical literature on "moral dumb-

[28] Ibid., 35.

[29] Recall Nietzsche, *GM* I, 13: "When out of the vengeful cunning of powerlessness the oppressed, downtrodden, violated say to themselves: 'let us be different from the evil ones, namely good! And good is what everyone is who does not do violence, who injures no one, who doesn't attack, who doesn't retaliate, who leaves vengeance to God, who keeps himself concealed, as we do, who avoids all evil, and in general demands very little of life, like us, the patient, humble, righteous'—it means, when listened to coldly, and without prejudice, actually nothing more than: 'we weak ones are simply weak; it is good if we do nothing *for which we are not strong enough*.'"

[30] Nietzsche, *D*, 99.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

founding," which finds that people will often remain attached to a moral judgment, even when all their reasons for it are defeated[31] (We will return to that literature at the end of the essay.)

If to individuate the meta-affect correctly, we need to take account of the cognitive judgment that informs it, then our picture would be complicated somewhat: we would have noncognitivism about basic affects, and a cognitivist view about meta-affects. On the other hand, it might suffice for causal explanation of behavior to individuate only the meta-affect of aversion or inclination toward the basic affect. In other words, what might matter for explaining the behavior of the Homeric Greek afflicted with the basic affect of aversion toward giving offense to dangerous enemies is that he feels a motivational push *against* acting on that basic affect, i.e., he feels aversion towards his basic aversion toward offending dangerous enemies. His cognitive beliefs may be causal triggers for the meta-affect, but we need only understand the meta-affect to understand his behavior.

Is that right? When the Homeric agent judges his aversion toward offending a stronger enemy as "cowardice," and feels "shame" or "guilt" about it, that judgment is a symptom of his basic affect of aversion toward offending dangerous enemies and his meta-affect of aversion toward the basic affect. But is it really explanatorily idle whether or not that meta-affect is one of "guilt" or "shame"? Take a slightly different Greek example. Recall that Oedipus, upon realizing he has killed his father and married his mother, gouges out his eyes, because he is overwhelmed with *shame* and so does not want to look into anyone's eyes again. Of course, Oedipus did not freely and intentionally choose to kill his father and marry his mother, so his overriding moral emotion is not one of guilt—but what if it were? Eliminating the possibility of human eye contact would not relieve the pain of *guilt*, since guilt, as usually construed, does not require an observer. The agent who experiences the meta-aversion (of "guilt") to the basic aversion of not marrying one's mother or killing one's father would believe he was *responsible* for having done this and thus blameworthy for his lapse of judgment. But, of course, Oedipus does not believe any of that, since he did not freely do what he did; he was fated to do so, which is why his meta-aversion is that of shame rather than guilt. So his meta-affective aversion toward his aversion toward killing his father and marrying his mother is of a particular kind: it is *ashamed aversion*, which can be blunted by eye-gauging, rather than *guilty aversion*, which could not be. It thus seems that *sometimes* the distinctively moral emotion is constituted in part by a particular cognitive content (for example, the one—whatever it is precisely—that separates shame from guilt), and so Nietzsche can not be a thoroughgoing noncognitivist about the affects underlying moral judgment, even if he is a noncognitivist about basic affects.

[31] Cf. Jonathan Haidt, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108 (2001): 814–34.

That conclusion would, however, explain why Nietzsche holds out the hope that attacking the falsity of the judgments that are the "grandparents" of the meta-affects could make a difference: "We have to *learn to think differently* [*umzulernen*]—in order at least, perhaps very late on, to attain even more: *to feel differently* [*umzufühlen*]."[32] To be sure, it could be that correcting such mistaken judgments could make a difference to even noncognitive affects, insofar as those judgments *happen to be* causally connected to the affects. But if such judgments are partly constitutive of the meta-affects, then the attack on the truth of the metaphysical presuppositions about agency (such as free will) that I have argued elsewhere is a key part of Nietzsche's critique of morality would be especially relevant.[33]

None of the preceding considerations affect my basic thesis, however: namely, that for Nietzsche, moral judgments are produced by affective responses. The complication we have noted is that drives give rise to feelings of inclination or aversion, but *not* to moral affects proper, the latter arising from the influence of culture (e.g., Homeric Greece versus Christian) which yield meta-affective responses to the basic affects of inclination and aversion. Moral judgments are *still* symptoms of the affects, but the affects now have two different sources: there is the affect of inclination or aversion toward particular actions or things or persons in the world, and then there are affective responses to those basic affects—where the latter are largely inculcated by the local culture and *perhaps* also individuated by their cognitive content. The latter observation also explains why Nietzsche expresses the optimistic view that revising our beliefs might actually lead to a revision of our feelings, i.e., those meta-affects or meta-feelings whose object is partly individuated by the cognitive judgment in question.

Someone might object, of course, that if moral judgments are indeed symptoms of *affects,* but some of these affects (the meta-ones) are either contingently or perhaps constitutively (in the case of, say, guilt versus shame) tied to beliefs, then moral judgments are not *only* the causal products of affective responses: they do depend on what people *believe,* not simply how they *feel.* That might suggest that Nietzsche's view of moral judgment would have to be tempered to say that moralities are symptoms of *affects, but not only affects.* I actually think Nietzsche has available a more ambitious response, since he also seems to hold the view that "belief fixation"—that is, the doxastic state in which an agent *takes a belief seriously enough that he will act on it*—is itself dependent on affective investment in the belief (think, for example, of his explanation of how a desire to punish motivates belief in free will). I will not be able to defend that view here, and it will suffice for the rest of the essay if the reader is

[32] Nietzsche, *D,* 103.
[33] Brian Leiter, *Nietzsche on Morality* (London: Routledge, 2002), 87–101.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

persuaded that moralities are symptomatic of affects, even if in some cases, they are symptomatic of affects *plus certain beliefs.*

IV. Affects, Drives, and the Doctrine of Types

How does the preceding picture of affects and drives, and their role in moral judgment, fit with what I have called Nietzsche's "Doctrine of Types,"[34] that is, his view that each person has some relatively stable set of psycho-physical traits that figure centrally in the explanation of his behaviors and values? The *psychological* component of a person's type consists of *drives,* as suggested by the famous passage on "great" philosophers in *Beyond Good and Evil,* which concludes that the "morals" of a philosopher "bear decided and decisive witness to *who he is*—which means, in what order of rank the innermost drives of his nature stand with respect to each other."[35] Yet, as we learned from *Daybreak,* drives stand in some relationship to both *basic* and *moral* affects,[36] and the latter reflect—given the account in *Daybreak*— the influence of parental and cultural conditioning.[37] Yet, as I have argued in previous work,[38] Nietzsche is not shy about ascribing *essential* aspects to humans, as when he refers to "the weakness of the weak . . .—I mean [their] *essence* [*Wesen*][39] or when he writes in 1886 that "assuming that one is a person, one *necessarily* [*nothwendig*] has the philosophy that belongs to that person."[40] A philosophy would only *necessarily* belong to a particular person if there were facts about that person—such as facts about his drives—that necessitated certain moral views, which, in turn, necessitated a certain philosophical position in their service. The impression that at least some part of the individual's psychic makeup is "hardwired" (as the contemporary metaphor has it) is reinforced by Nietzsche's Lamarckianism, that is, his embrace of the view (widely accepted at the time) that culturally acquired traits of the parents can then be inherited by the offspring.[41] The irony, from the standpoint of the early twenty-first century, is that while we now know Lamarck was wrong, the *heritability* of characterological traits—such as a propensity towards violence (or "aggressive antisocial behavior" as psychologists usually call it)—is, in fact, well established in

[34] Leiter, Nietzsche on Morality, 8–10,
[35] Nietzsche, *BGE,* 6.
[36] Nietzsche, *D,* 38,109.
[37] See, e.g., Nietzsche, *D,* 104: "[W]e arrive [at our evaluation, which causes actions] as *children,* and rarely learn to change our view; most of us are our whole lives long the fools of the way we acquired in childhood of judging our neighbors (their minds, rank, morality, whether they are exemplary or reprehensible) and of finding it necessary to pay homage to their evaluations."
[38] Leiter, *Nietzsche on Morality,* 26.
[39] Nietzsche, *GM,* I:13.
[40] Nietzsche, *GS,* "Pref.," 2.
[41] Cf. Nietzsche, *GS,* 143; *BGE,* 213.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

behavioral genetics.[42,43] Of course, such traits naturally have a certain culturally specific component that might seem to lend support to Lamarckianism, yet they are not actually *inherited,* since the trait itself has no gene. But such traits can be *heritable,* and thus mimick the Lamarckian result, because of the interplay of genetically determined traits and cultural milieu. Let me explain.

Here is a simple example:[44] a gene for having a large nose can be inherited, but in a cultural environment where large noses are objects of ridicule (such that their possessors end up being quite shy), shyness, as a character trait, will turn out to be *heritable,* not because of a "shyness gene" but because of a "large nose gene" conjoined with a stable cultural environment which interacts with the inherited phylogenetic features of human organisms to produce the psychological attribute. If, like Lamarck, you do not know that genes are the mechanism by which evolution by natural selection operates, one might well think that culturally acquired traits can be inherited.

Nietzsche knew as little about the difference between inheritance and heritability as Lamarck, but that is not the point here. The point is that it is not crazy to think of even character traits as among the type-facts about a person, at least against the background of a particular cultural environment. So, to return to our earlier example, if the drive to avoid dangerous enemies is inherited—in the sense that an aversion towards such enemies is "hardwired" or part of a person's type—and the cultural environment is more or less stable in its Christian values, then it will certainly *appear* that "humility" is a "hardwired" part of the slavish type of person.

Of course, the crucial idea here is that Nietzsche thinks certain affective responses are either inherited or heritable, such that they are facts about an agent over which the agent has no control. In the case of drives that give rise to basic affects, they can be simply inheritable; in the case of meta-affects, as long as they are heritable, the situation will be the same for the agent. I think it is clear that something like this is Nietzsche's

[42] Joshua Knobe and Brian Leiter, "The Case for Nietzschean Moral Psychology," in *Nietzsche and Morality,* ed. Brian Leiter and Neil Sinhababu (Oxford: Oxford University Press, 2007).

[43] Typically, these studies are conducted either by looking at twins (comparing monozygotic to dizygotic) or by looking at adopted children. One study, for example, using 1,523 pairs of twins found a heritability of aggressive antisocial behavior of 70 percent (Thalia Eley, Paul Lichtenstein, and Jim Stevenson, "Sex Differences in the Etiology of Aggressive and Nonaggressive Antisocial Behavior: Results from Two Twin Studies," *Child Development* 70 [1999]: 155–68.) Other studies yield percentages that are lower but still surprisingly high—for example, 60 percent (C. Edelbrock, R. D. Rende, R. Plomin, and L. A. Thompson, "A Twin Study of Competence and Problem Behavior in Childhood and Early Adolescence," *Journal of Child Psychology and Psychiatry* 36 [1995]: 775–85, and 49 percent (K. Deater-Deckard, and R. Plomin, "An Adoption Study of the Etiology of Teacher Reports of Externalizing Problems in Middle Childhood," *Child Development* 70 [1999]: 144–54.) These huge effect sizes cannot plausibly be ascribed to experimental artifacts or measurement error, as opposed to genetic influences.

[44] And a memorable one, which I owe to Jesse Prinz.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

view—not a surprising view in the nineteenth century, of course—but it is worth emphasizing that it is also not, in light of subsequent research, an implausible view.[45]

V. Is It True That Moralities Are Just a Sign-Language of the Affects?

We have now given a systematic exposition of Nietzsche's idea that moral judgments are symptoms or sign-languages of the affects, in the sense that moral judgments are caused by affects and meta-affects, which

[45] A different kind of question can also be raised about Nietzsche's picture of moral psychology: namely, why should we think drives exist in the first place? In many passages, after all, Nietzsche notes how ignorant persons are about the drives that constitute them: "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," as he puts it in *Daybreak* (119). Since drives are not identified directly by their conscious, qualitative feel, the grounds for positing their presence must amount to something like an inference to the best explanation of observable behavior and conscious feelings. Yet Nietzsche is notoriously promiscuous with his ontology of drives; in *Daybreak* alone he refers to "the drive to restfulness, or the fear of disgrace and other evil consequences" (*D* 109); the hunger drive (*D* 119); "moral . . . drives" (*D* 119); "our drive to tenderness or humorousness or adventurousness or . . . our desire [*Verlangen*] for music and mountains" (*D* 119); "the drive to praise or blame" (*D* 140); the drive to feel sorrow or sadness (when triggered by music) (*D* 142); and "the drive to attachment and care for others (the 'sympathetic affection')" (*D* 143). On top of that, Nietzsche also claims that drives vary in their degree of *strength* or *vehemence* (*Heftigkeit*) (*D* 109)—that is, how badly they need to be satisfied—and that this quality of drives is important to their role in explaining moral judgment and action (cf. *D* 119).

How are we to appraise Nietzsche's promiscuous ontology of drives? There is not even the pretense of demonstrating the explanatory need for some of the drives he posits, many of which are on a par with Moliere's doctor's explanation that opium makes one sleep because of its dormitive power! It is worth bearing in mind, however, that the interest of Nietzsche's hypothesis about the role of drives in moral judgment does *not* depend on any specific hypothesis about particular drives; it turns, rather, on the correctness of the basic model of the mind. We may view some of the specific claims simply as "placeholders" for a more adequate psychology, and it is reasonable to think some of the specific drives Nietzsche posits are explanatorily otiose and would drop out of a more systematic account. (I am indebted here to Roger Eichorn.)

Drives for Nietzsche are (per Katsafanas's formulation) "dispositions that induce affective orientation in the agent." These affective orientations structure how the world appears to us evaluatively: they influence, for example, "perceptual salience," the features of a situation that come to the fore for the agent (because the drive focuses attention on them). (Cf. *D* 119. Humeans also have a similar view: see, e.g., Neil Sinhababu, "The Humean Theory of Motivation Reformulated and Defended," *Philosophical Review* 118 [2009]: 465–500; 469–70, discussing the "attention-direction" aspect of desire.) Katsafanas also argues that Nietzschean drives share two features of drives on the Freudian view (unsurprisingly given Freud's interest in Nietzsche). First, drives have a kind of constancy that particular desires do not. The music you desired to listen to in your twenties may no longer appeal in your forties; but the hunger drive keeps coming back whether you are twenty or forty. Second, drives do not depend on an external stimulus to be aroused. External stimuli can give rise to a desire to eat or to have sex, to be sure, but those same desires can simply arise in the absence of any stimuli. It is particularly useful to distinguish, as Freud does (33), between the *Ziel* (aim) of the drive (e.g., sex, eating) and the *Objekt* of the drive (e.g., this woman, this bit of food). Insofar as a drive is aroused not by an external stimulus, it will then seek out an object for its realization—and in so doing impose a "valuation" on the object.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

are the joint product of nature and culture, as it were. This is, I have argued, Nietzsche's view, but the more pressing philosophical question is: Why should we accept it?

The idea that moral judgments are caused by affective responses has three kinds of evidence in its favor: first, sincere moral judgments—not the lip-service to morality that is the stuff of so much public life—motivate action (or at least produce what we earlier called *motivational oomph*); second, it is a view of moral judgments compatible with one plausible metaphysics of morality, namely, antirealism (the view that there are no objective moral facts), a view Nietzsche himself explicitly endorses; and third, it squares best with existing empirical research on the psychology of moral judgment and moral disagreement. The latter two considerations are not decisive, of course, since it might turn out that Nietzsche held false views, if it happened that moral realism were true or that empirical psychology gave us a different account of moral judgment and disagreement. But if there is good textual evidence that Nietzsche held a view that also seems to be true, then that seems an additional reason of interpretive charity to ascribe it to him. Let us take these considerations up in turn.

First: sincere moral judgments motivate action or, at least, produce motivational oomph. If I genuinely judge that X is morally wrong or Y is morally right, I feel motivated, respectively, to prevent X or bring about Y, even if I fail to do either—cowardice, weakness of will, timidity and so on may all intervene to prevent people from acting on their moral judgments. But what is striking is that those who sincerely affirm a moral judgment at least feel a motivational tug in its direction. That fact has long been thought to be one of the strongest reasons in favor of antirealist views of moral judgment. More precisely, it has been thought to count in favor of *noncognitivist* views of moral judgments, according to which such judgments simply express motivationally effective mental states, such as desires and feelings, rather than mental states like belief that only represent aspects of the world. Nietzsche, as I have argued previously,[46] does not have a considered view about the semantics of moral judgment, which is hardly surprising given that the issue was not a live one prior to the twentieth century, and given that it is an issue that turns on such subtle linguistic considerations that many philosophers have now gravitated toward "hybrid" views in which the correct analysis of the semantics is thought to incorporate both cognitivist and non-cognivist elements.[47] I suggest we prescind from these issues, both for interpretive reasons (Nietzsche has no real view on them) and philosophical ones. On the philosophical front, remember that the primary consideration thought to militate

[46] Brian Leiter, "Nietzsche's Metaethics," *European Journal of Philosophy* 8 (2000): 277–97.
[47] For doubts about these, see Mark Schroeder, "Hybrid Expressivism: Virtues and Vices," *Ethics* 119 (2009): 257–309.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

against a noncognitivist interpretation of the semantics has always been the Frege-Geach problem, the problem of how to explain the truth-preserving properties of inferences involving moral propositions embedded in the antecedents of conditionals (e.g., "If stealing is wrong, then it is wrong to encourage John to steal"; but "stealing is wrong"; so "it is wrong to encourage John to steal"). But it would be mad to let our metaphysics—our most plausible account of what really exists—be driven by linguistic practices: Why let the semantic tail wag the metaphysical dog? As Crispin Wright observed twenty years ago, if metaphysical antirealism about moral facts conjoined with noncognitivism about the semantics of moral judgments had "absolutely no prospect of a satisfactory construal of conditionals with moral antecedents that could hardly be decisive. Rather, whatever case there was for [this kind of antirealist view] would become potentially revisionary of our ordinary and moral linguistic practice. . . ."[48] If we do not want to "bite the bullet" on such "radical revisionism" (as Wright aptly calls it), then there remain two main options on the semantic front: first, there are highly technical noncognitivist solutions to the Frege-Geach challenge, like Allan Gibbard's; and second, we can adopt a minimalist approach to truth, such that the propriety and intelligibility of certain assertoric idioms in evaluative language are enough to warrant cognitivism, with the issue between moral realists and antirealists located elsewhere (for example, in the conception of objectivity).[49]

Even if we bracket the semantics, and view the thesis that moral judgments are a "sign-language" of the affects as primarily a causal hypothesis—which is how I have suggested we take it—then the considerations just adduced that count in favor of noncognitivist semantics will also count in favor of Nietzsche's actual hypothesis. For if agents making sincere moral judgments are motivated to act in accordance with those judgments (whether or not they follow through on the motivation), then it counts in favor of any hypothesis about the causes of moral judgments that it should explain why such judgments produce motivational oomph. But the hypothesis that moral judgments are caused by affective responses is precisely that kind of hypothesis, since everyone agrees that affective or emotional responses put causal pressure on action. So the first reason to believe that moral judgments really are just "sign-languages of the affects" is that it explains why moral judgments motivate their makers, that is, cause them to feel inclined to or averted from certain actions, even if they do not actually act upon those feelings.

Second: any acceptable view of moral judgment ought to be consistent with a plausible account of the metaphysical status of moral facts, and

[48] Crispin Wright, "Realism, Antirealism, Irrealism, Quasi-Realism," *Midwest Studies in Philosophy*, ed. P. French et al. (Notre Dame, IN: University of Notre Dame Press, 1988), 31.

[49] Crispin Wright, *Truth and Objectivity* (Cambridge, MA: Harvard University Press, 1992) is the *locus classicus* for this kind of view.

any plausible view of moral judgment ascribed to Nietzsche should, in particular, be compatible with his own view about the metaphysical status of moral facts. I have argued elsewhere[50] that Nietzsche's metaphysical view, following the Presocratics, is clear: there are no objective facts about what is morally right and wrong. Nature is, as he says in *The Gay Science,* "always value-less, but has been given value at some time, as a present—and it was we who gave and bestowed it." Such a metaphysical picture fits neatly with the idea that our judgments about moral value are not responsive to the preexisting evaluative features of the world, but rather are caused by noncognitive affective states of the judger.[51]

Let us look at two illustrative statements of Nietzsche's view from *Daybreak,* his first "mature" work as it were. Early on in book I[52], he compares the way "man has ascribed to all that exists a connection with morality [*Moral*] and laid an *ethical significance* [*ethische Bedeutung*] on the world's back" to the earlier, and now discredited, "belief in the masculinity or femininity of the sun." Later, in book II of the same work[53], he compares the way "wise and noble men still believe in the 'moral significance [*sittliche Bedeutung*] of existence'" to the way they previously "believed in the music of the spheres," which is "no longer . . . audible to them." (The switch to *sittliche Bedeutung* probably reflects the fact that the theme of the "morality of custom [*Sitte*]"—i.e., *Sittlichkeit*— was introduced in book I.)

Passages like these invite both a metaphysical and an epistemological interpretation (I put aside the possible semantic reading, for the reasons already given). On the *metaphysical* reading, Nietzsche is saying that *there do not exist objective moral properties in the world,* just as there do not exist gendered properties of the sun, or musical properties of the heavenly bodies. To be sure, at one time, humans "perceived" the sun and the heavenly bodies as having gendered and musical properties, but those entities did not really (objectively) have them. Nietzsche's view thus seems to be a kind of *projectivism,* in particular, an instance of what Peter Kail (in discussing Hume) calls *explanatory projection* according to which "some feature of our mentality explains how the world appears to us." In a case of *explanatory projection,* "[t]he thinker is not responsive to the world in a way whereby their beliefs, concepts and experience reflect their object. In a slogan: explanatory projection is non-detective [or non-detecting] expla-

[50] Leiter, *Nietzsche on Morality,* 146–150; Brian Leiter, "Moral Skepticism and Moral Disagreement in Nietzsche," in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, volume 9* (forthcoming 2014).

[51] I concede it is possible that someone might claim that affective responses are epistemically reliable ways of tracking the truth about what morality requires. That is plainly not Nietzsche's own view, and, given the overwhelming evidence of the ways in which emotional responses are epistemically unreliable in so many other contexts, it would be surprising, indeed, were they to turn out to be epistemically superior in this domain.

[52] Nietzsche, *D,* 3.

[53] Ibid., 100.

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

nation"[54] On the e*pistemological* interpretation, which I think Nietzsche probably also intends, our purported *knowledge* of objective moral facts is an illusion, an unsurprising upshot of the metaphysical reading. As Nietzsche puts it, once again in *Daybreak:* "it is a prejudice of the learned that *we now know better* than any other age . . . what is good and evil."[55]

Now the ascription of moral antirealism in Nietzsche's case is complicated by the fact that he does not seem to make the mistake of some earlier twentieth-century antirealists, like A. J. Ayer and Charles Stevenson, who simply assumed that the metaphysical status of moral norms was different in kind from the status of epistemic norms. After all, as we have already seen, it is nature "that is value-less," and *all* value (*Werthe*) is "bestowed" by humans onto this value-free nature. But judgments about what we *ought* to believe in light of the evidence also depend on *values*—"epistemic norms" as we usually say nowadays—and it is hard to see why those values should be exempt from Nietzschean antirealism. All knowing is, after all, as Nietzsche argues in the famous *Genealogy* passage on perspectivism, animated by *affects* (or interests),[56] but if those are, as I have argued noncognitive, then it is hard to see how they could be deemed epistemically reliable or special. The point is even more explicit in the late discussion of perspectivism in *The Gay Science* (in book V, added in 1886), when Nietzsche says that "we 'know' (or believe or imagine) exactly as much as is *useful* to the interests of the human herd, to the species: and even what is here called 'usefulness' is finally also just a belief, an imaginary construct . . ."[57] These passages are important, in this context, not because they express doubts about truth or its existence, but because their target is explicitly *epistemic,* suggesting that norms of epistemic warrant answer to interests and affects that, themselves, have no independent standing as reliable trackers of the truth.

The conjunction of an apparent global antirealism about value and specifically moral antirealism can appear perplexing, and has confused some Anglophone philosophers.[58] On the one hand, Nietzsche seems to deny the metaphysical objectivity of values on the basis of broadly naturalistic considerations, considerations that themselves suppose certain epistemic norms, the norms in virtue of which we deem nature to have the nonmoral characteristics it has. Yet Nietzsche seems to have no reason

[54] The contrast is with what Kail calls "feature projection," which he also finds in Hume, and which is central to the Freudian concept of projection: "In features projection, features of our mentality become represented as features of some other object (I project my hate in thinking that someone else hates me)." (Peter Kail, *Projection and Realism in Hume's Philosophy* [Oxford: Oxford University Press, 2007], xxix–xxx.)

[55] Nietzsche, *D,* 2.

[56] Nietzsche, *GM,* III, 12.

[57] Nietzsche, *GS,* 354. *[W]ir 'wissen' (oder glauben oder bilden uns ein) gerade so viel als es im Interesse der Menschen-Heerde, der Gattung,* nützlich sein mag: *und selbst, was hier "Nützlichkeit' gennant wird, ist zulezt auch nur ein Glaube, eind Einbuildung. . . ."*

[58] Hilary Putnam, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2004), on the fact-value distinction is a good example.

to exempt epistemic values from his value skepticism. How then should his position be understood?

There is no reason, in my view, to think Nietzsche had a clear handle on this philosophical issue. Like a lot of intellectuals self-taught in philosophy, he tended to make a mess of certain philosophical problems, especially those related to knowledge and truth. Yet the problem his metanormative skepticism presents is one that has become familiar in the second half of the twentieth century, one that figures in the work of W. V. O. Quine and John McDowell, among other leading philosophers of the past half-century. Let me propose, then, a resolution that may or may not have been one Nietzsche had in mind, but which makes good philosophical sense of his position (and, indeed, seems to me a sensible position).

Suppose one thinks, with Nietzsche (and arguably others), that no values, moral or epistemic, can claim any special metaphysical standing, that nature is really "value-less," meaning, among other things, that nature does not settle what moral values one should accept or what *epistemic* values one should accept. Various lessons from twentieth-century philosophy certainly push in that direction. First, from the famous Duhem-Quine thesis about the underdetermination of scientific theories by evidence, we know that there are not even any scientific hypotheses that are epistemically obligatory, in the sense of required by the norms of logic and evidence.[59] This is because any recalcitrant evidence elicited in a test of a hypothesis is compatible with the hypothesis as long as we are willing to give up the background assumptions such a test requires. In choosing among competing hypotheses and background assumptions, we must always fall back on evaluative considerations that "nature" does not adjudicate among, considerations such as theoretical simplicity, methodological conservatism, and consilience.[60] Second, unless there were a plausible *substantive* conception of rationality (there does not appear to be one, alas), then rationality, including any internalist norm of epistemic warrant, is itself instrumental, imposing normative constraints only on the means chosen to realize our ends, whatever they may happen to be. Thus, even norms for belief are hostage to ultimate ends, and so particular beliefs are unwarranted (that is, irrational) only relative to the believ-

[59] Pierre Duhem, *La Theorie Physique:Son Object et sa Structure,* (Paris: Marcel Riviera et Cie, 1914); W. V. O. Quine, "On Empirically Equivalent Systems of the World," *Erkenntnis* 9 (1975): 313–28 and "Three Indeterminacies," in *Perspectives on Quine,* ed., R. B. Barrett and R. F. Gibson (Cambridge, MA: Blackwell, 1990).

I acknowledge that the inveterate dogmatic realist may think this is merely an *epistemic* point, not a metaphysical one: there could still be *real epistemic values,* after all, we just do not know what they are or how to apply them. That is a logically possible position, but I am with Quine in thinking that if the *actual successful sciences* do not disclose such epistemic values, then it is dubious that reality demands any particular set of them.

[60] Cf. W. V. O. Quine and Joseph Ullian, *The Web of Belief,* 2d ed. (New York: Random House, 1978).

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

er's ends.[61] That would hardly be surprising to Nietzsche, who clearly appreciated the extent to which theoretical questions were driven by practical ends and interests.[62] Why, then, think there is a special metaphysical problem for *moral* values?

The idea of a "special" problem for metaphysical realism about moral values is ambiguous, however, between a problem *in kind* as opposed to one *of degree.* One upshot of the preceding considerations is that there may be no metaphysical difference *in kind* between moral and epistemic values, but that is still compatible with a radical difference in degree between them, and I think that is the key to making sense of Nietzsche's view. Let us call a "Global Humean" about epistemic values someone who notices that creatures like us generally converge on epistemic norms because those norms do so well at meeting widely shared human needs and interests. Consider epistemic norms like the following: treat as prima facie veridical perceptual evidence, honor logical inferences, and employ the inductive method in empirical inquiry. These epistemic norms seem to facilitate successful navigation of the world and prediction of the future course of experience. Something like this, I suspect, was Hume's own view, though unlike Hume, Nietzsche does not think natural dispositions converge as well in the ethical case. That would explain why the great insight Nietzsche attributes to the Sophists concerns "the multiplicity (the geographical relativity) of the moral value judgments [*Moralischen Werthurtheile*]"[63]—not *all* value judgments, in other words, but the distinctively moral ones.

Third and finally: Nietzsche's view that moralities are a sign-language of the affects in something like the sense I have articulated here fits well with the empirical evidence about moral judgment and motivation. We should acknowledge, of course, that the evidence is ambiguous, but it certainly tends in the Nietzschean direction. Recall, for example, Nietzsche's claim, discussed earlier, that "moral feelings" (*Moralische Gefühle*) are inculcated when "children observe in adults inclinations for and aversions to certain actions and, as born apes, *imitate* these inclinations [*Neighungen*] and aversions [*Abneighungen*]; in later life they find themselves full of these acquired and well-exercised affects [*Affekten*] and consider it only decent to try to account for and justify them."[64] That point fits nicely with Jonathan Haidt's famous work on the "social intuitionist" model of moral judgment according to which in most ordinary situations, moral judgments are produced by emotional or affective responses, the reasons adduced in their support being post hoc: they do not explain the judgment, as evidenced by the resilience of the judgment even in the face of the defeat of the prof-

[61] Peter Railton, "Facts and Values," *Philosophical Topics* 14 (1986): 5–31.
[62] Nietzsche, *BGE,* 3–9.
[63] Nietzsche,*WP,* 428.
[64] Nietzsche, *D,* 34.

fered reasons.[65] Haidt, himself, is somewhat confused about the import of these empirical findings, for, contrary to Haidt, they state no dispute with a philosophical rationalist about moral judgment like Kant, since Kant is not committed to *either* the claim that most people actually arrive at their moral judgments through the exercise of practical reason *or* the claim that most people arrive at moral judgments that can be justified, even after the fact, by the correct exercise of practical reason. The Kantian rationalist is committed, I take it, only to the claim that rational agents can, in principle, revise their moral judgments in light of practical reason, but nothing in Haidt's research rules out that possibility—indeed, he acknowledges that sometimes reasoning can result in a revision of moral judgments. But Nietzsche needs for his purposes only the descriptive thesis—that affective or emotional responses ordinarily determine moral judgment—since he has independent arguments for skepticism about practical reason against the moral rationalist like Kant that do not depend on the actual causal process by which people ordinarily arrive at moral judgments.

Recent empirical work on moral psychology also lends support to Nietzsche's view of moral judgments as sign-languages of the affects. As a recent literature survey notes, individuals "with selective deficits in emotional processing" due to disease or injury to the brain, render different moral judgments about hypothetical situations like the Trolley cases, than most emotionally normal subjects,[66] suggesting that the affective responses are causes of the moral judgments. Psychologist and philosopher Joshua Greene[67] has argued that emotional responses loom larger in deontological than consequentialist moral judgments, the latter demanding more "controlled cognition," but in more recent work even Greene has acknowledged that "affect supplies the primary motivation to view harm as a bad thing" in the first place, so that even consequentialist reasoning has "an affective basis."[68]

In a recent review of the empirical literature, Timothy Schroeder, Adina Roskies, and Shaun Nichols found that the view they dub

[65] Haidt, "The Emotional Dog and Its Rational Tail."

[66] Fiery Cushman, Liane Young, and Joshua D. Greene, "Multi-System Moral Psychology," in *The Moral Psychology Handbook,* ed., John M. Doris and the Moral Psychology Research Group, (Oxford: Oxford University Press, 2010), 53–54.

[67] Joshua Greene, "The Secret Joke of Kant's Soul," in *Moral Psychology, Volume 3,* ed. W. Sinnott-Armstrong (Cambridge, MA: MIT Press, 2007): 35–117.

[68] Cushman et al., "Multi-System Moral Psychology," 54; 62.

More precisely, "affect supplies the primary motivation to regard harm as bad. Once this primary motivation is supplied, reasoning proceeds in a currency-like manner ["currency emotions are designed to participate in the process of practical reasoning"]" (Cushman et al., "Multi-System Moral Psychology," 63.) "[A]larm-bell emotions are designed to circumvent reasoning" (*id.* at 62) and, arguably, this is "the origin of the welfare principle," namely, "Parkinson disease appears to show that intrinsic desires are necessary to the production of motivation in normal human beings, and this would seem to put serious pressure on the cognitivist position" (93).

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press

"sentimentalism"—namely, the view that "the emotions typically play a key causal role in motivating moral behavior"—is well supported by the "Evidence from psychology and neuroscience," and that while "motivation derived [exclusively] from higher cognitive centers independently of desire is possible . . . the only known model of it is pathological" involving Tourette Syndrome.[69] Such empirical findings do not rule out the possibility that moral judgments are not causal products of the affects, of course, but they suggest that the evidential burden must be borne by views that deny that causal role to affective responses.

Nietzsche's account of moral judgments as sign-languages of the affects seems, in short, to have the empirical evidence on its side, and this is not the first time a Nietzschean hypothesis has turned out to win support from subsequent empirical psychology, as I have argued elsewhere.[70] But that fact raises a question of its own, which I would like to address in concluding. Any reader of Nietzsche knows that he is not primarily concerned to report the findings of psychological "research" or to establish, through conventional methods of argumentation and the mustering of evidence, the truth of particular empirical hypotheses. His work is suffused with psychological and empirical claims, to be sure, but his aims are always much more polemical: to transform the consciousness of at least some readers about the morality they take for granted, and thus, at the same time, to change their affective orientation toward their lives. Nietzsche's psychological claims are always subservient to these rhetorical aims, but this does not alter the fact that he makes psychological claims, ones that admit of empirical study and confirmation, and many of which may, as recent research suggests, be true. But, how one might wonder, could Nietzsche have been so right about so much of moral psychology without employing the methods of contemporary empirical psychology?

To start, we must remember that the best account of the psychology of human agency, moral and otherwise, is on a clear continuum with ordinary "folk" psychological explanations of behavior, in a way that contemporary physical science is *not* derivable from "folk" physics, that is, from the ordinary categories we use to make sense of our observations about the physical world around us. Atoms, molecules, and invisible forces do not play an obvious role in my observation that if I drop the sofa on my toe, it will hurt, but contemporary empirical psychology avails itself of basically the same ontology—beliefs, desires, traits, bodily movements—that are the very stuff of unsystematic folk psychology.

To be sure, empirical psychology has plainly evolved methods for testing and confirming hypotheses that were not in use in the nineteenth

[69] Timothy Schroeder, Adina L. Roskies, and Shaun Nichols, "Moral Motivation," in *The Moral Psychology Handbook,* ed. Doris et al., 77; 98; 94.

[70] Brian Leiter, "Nietzsche's Theory of the Will," *Philosopher's Imprint* 7 (2007): 1–15; Knobe and Leiter, "The Case for Nietzschean Moral Psychology."

century, but that does not mean Nietzsche lacked *evidence* on which to base his speculative moral psychology. Speculation by geniuses based on limited evidence has always played a major role in scientific progress, and it should hardly be surprising if psychology were any different. Nietzsche's own evidence appears to have been of three primary kinds. First, there were his own observations, both introspective and focused on the behavior of others. Second, Nietzsche was an avid consumer of the psychological observations recorded by others, in a wide array of historical, literary, and philosophical texts over long periods of time, observations that, in some respects, tended to reinforce each other. (Consider, for example, the realism about human motivations detailed by Thucydides in antiquity and, in the modern era, in the aphorisms of La Rochefoucauld, both authors whom Nietzsche admired). Finally, there was Nietzsche's extensive reading about contemporaneous scientific developments in the 1850s and 1860s, most of which—even if amateurish or simply wrong by today's standards—did represent systematic attempts to bring scientific methods to bear on the study of human beings and many of which, in at least some of their broad outlines, have been vindicated by subsequent developments. By the standards of contemporary methods in the human sciences, we would not deem insights based on this evidence to be well confirmed, but that certainly does not mean such evidence is not, in the hands of a genius like Nietzsche, adequate for insights that survive scrutiny by our contemporary methods. This is precisely one of the reasons why Nietzsche, like Hume, is a great *speculative* naturalist in the history of philosophy: with unsystematic data and methods he could nonetheless arrive at hypotheses—including the hypothesis that moral judgments are sign-languages of the affects—that turn out to be supported by the more systematic data and methods the scientific study of human beings relies on today.

Law, Philosophy, & Human Values, University of Chicago

https://doi.org/10.1017/S0265052513000113 Published online by Cambridge University Press