

Nietzsche and Moral Psychology

DANIEL TELECH AND BRIAN LEITER

A remarkable number of Nietzsche's substantive moral psychological views have been borne out by evidence from the empirical sciences. While a priori officiating on the nature and value of psychological phenomena (e.g., belief, volition, desire) continues unchecked in some circles, Nietzsche both anticipated and would welcome the wealth of empirical research on the natural bases of our evaluative attitudes. Indeed, Nietzsche holds that "[a]ll credibility, good conscience, and evidence of truth first come from the senses" (BGE 134). Understanding that empirical evidence was likely to challenge many of our deeply held self-conceptions, it is not surprising that Nietzsche (BGE 23) calls for individuals "daring to travel" the path of psychological inquiry. Partly owing to its capacity to uproot our stultifying traditional moral prejudices, "*psychology*," according to Nietzsche, "is the path to the fundamental problems" – accordingly, his hope is that psychology "be recognized as queen of the sciences."

The aim of this chapter is to introduce readers to Nietzsche's promise for empirically informed philosophical psychology by attending to four of his claims: (1) that moral responses are products of the affects; (2) that each person has a relatively stable psycho-physiological constitution that qualifies him or her as a "type"; (3) that conscious acts of willing are frequently epiphenomena of the real causal mechanisms of action; and (4) in spite of an absence of volitional freedom, self-control can be usefully understood on a "strength-model" of motivational resources. We conclude with a brief discussion of how Nietzsche, without employing the contemporary methods of empirical psychology, could nonetheless be such a prescient moral psychologist.¹

7.1 Moralities are Symptoms of the Affects

In opposition to Socratic moral intellectualism and the convenient Kantian "discovery" of moral faculties (BGE 11), Nietzsche argues that our moral commitments are causal products of the affects. As he puts it (BGE 187), "morality is just a *sign language of the affects!*"² Claims to the effect

that moral judgments are “symptoms” or “sign-languages” of drives and affects abound in Nietzsche’s corpus (see e.g., D 119; TI “Problem”: 2; GM “Pref”: 2; WP: 258).³ Nietzsche’s idioms of “symptoms” and “sign-languages” should be understood causally. To say that sweating is a symptom of a viral infection is to say that the symptom is the effect of the virus, and additionally, that the symptom provides us with inferential evidence for the existence of the virus (cf. Leiter 2013, 239). To hold that moral responses are symptoms of affects, then, is to say that moral responses are caused by, and reveal the existence of, certain affective states. In identifying moral response with affective phenomena, Nietzsche subscribes to a version of moral sentimentalism.⁴

We are disposed to have certain affective responses, on Nietzsche’s view, due to the organization of our *drives*.⁵ Drives are dispositions that structure our affective orientation and influence the salience of certain features in our environment (Katsafanas 2013, 740).⁶ Since affects are *essentially* valenced – they are states of inclination *to* or *from* – the motivational force of moral response is well accounted for on Nietzsche’s view. While affects are primarily noncognitive states, and so, individuated by the way they feel, that is, their phenomenal characters (D 34), they *do* have intentional objects, and so might be construed as states of “feelings toward” objects (cf. Goldie 2002, 19).

Affects are only *primarily* noncognitive because Nietzsche holds a two-level model of affective response, one level of which is sometimes individuated by reference to phenomenal character *and* propositional attitudes. While “basic affects” are wholly noncognitive states, we often display inclinations to and aversions from our basic affects, and these “*meta*-affects” may involve propositional attitudes. As Nietzsche claims,

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*... In itself it has, *like every drive*, neither this moral character nor any moral character at all. (D 38)

Drives are morally undetermined, so the basic affects they generate will not amount to moral emotions until their bearer takes a meta-affective stance (usually culturally shaped, and often involving beliefs) toward the basic affect – in this case, aversion to dangerous enemies. Given his beliefs about the nature and significance of courage, the ancient Greek’s basic aversion toward his enemy, for example, generates the affect of shame in him, yet the *same* basic affect gives rise to something pleasant like humility in the Christian.

On Nietzsche’s two-level view, then, moral affects are apt to be modified by cognition, yet they are nonetheless primarily noncognitive. Noncognitivism about moral responses earns support from a variety of sources. Haidt’s research on “harmless taboo violations,” for example, suggests that our practices of moral condemnation are ill-explained by considerations regarding harm. Rather, deep-rooted affects like disgust play an important explanatory role for many of our moral judgments (Haidt 2012, 26). Additionally, Greene’s (2007) work on moral motivation suggests that deontological reasoning is, *pace* the Kantian, especially influenced by affective response,⁷ and his more recent coauthored research (Cushman, Young, and Greene 2010, 53–4) suggests that *in both deontological and consequentialist reasoning*, “affect supplies the primary motivation to view harm as a bad thing.” The work of Cushman *et al.* also strongly suggests that emotional deficits significantly inhibit the ability of subjects to arrive at (what are widely considered to be) morally appropriate judgments when confronted with traditional thought-experiments. Additionally, research on criminal populations (Blair 1995) suggests that the inability of psychopaths to distinguish between moral and conventional transgressions is due in large part to defects in a mechanism responsible for aversion to intraspecies aggression (for discussion, see Nichols 2004, 12–16).⁸

Developmental psychology also favors noncognitivism, since it suggests that moral education initially proceeds via affective contagion and mimicry (Hoffman 2000, 36–9). Nietzsche's own view is that, “[m]oral 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). We not only make such post-hoc justifications to “render more respectable” our deepest beliefs, but there also exists evidence (Haidt 2001) that such “justifying” judgments are regularly insensitive to countervailing evidence.⁹

Deep-rooted though they may be, affective responses and moral judgments *are* alterable. They better be: Nietzsche is a revisionist about both of these (as we might expect given his overarching project of “revaluing all values”; cf. Reginster 2006, 148–9). Indeed, central to Nietzsche's project is the conviction that “[w]e have to *learn to think differently* – in order at least perhaps, very late on, to attain even more: *to feel differently*” (D 103). It is accordingly a virtue of his view that, while basic affects are individuated by phenomenal character alone, meta-affects (and the evaluative judgments they undergird) *can* be gradually transmuted, upon the rejection of “life-denying” beliefs. (For more recent proposals that we alter our inherited moral emotions (e.g., “reactive attitudes”) by revising our metaphysical views, see Pereboom 2001, 187–210; Sommers 2007.) The possibility of radical attitudinal revision, however, will depend upon still other facts about the individual.

7.2 The Doctrine of Types

Moral judgments are products of affects on Nietzsche's view, but the latter are in turn causally dependent upon more fundamental features of the individual. Nietzsche accepts a “Doctrine of Types” (Leiter 2015, 6), according to which,

Each person has a fixed psycho-physical constitution, which defines him as a particular *type* of person.

“Type-facts” consist in facts about the individual's physiology and unconscious drives, and for each person, there is some set of such facts that constitute him or her as a given *type* (for details, see Leiter 2015, 6–8; Leiter 2001, 294). Although such facts display a certain kind of *fixity*, they are not immutable: they can vary significantly in strength over time. Nietzsche's Doctrine of Types is to be distinguished from nineteenth-century vulgar biological determinism, since Nietzsche's view is about the causal *primacy* of (unconscious) psycho-physiological states, which does not amount to the causal sufficiency of such states (cf. Leiter 2015, 72–81). Not only are type-facts mutable, a person's type-facts at any given time do not completely determine her behavior; that is, though constitutive of *who one is*, type-facts stand in a non-necessitating relationship to one's behavior.¹⁰ This is not to deny that one's behavior is entirely necessitated. Rather, Nietzsche compares the necessity of human action to the “inevitability of fruits borne on the tree” (GM “Pref”: 2). The inevitability of the course of the tree's development is of course compatible with the fact that its seeds are alone insufficient for the existence of a mature tree. As Knobe and Leiter (2007, 90) elucidate Nietzsche's point,

Think of some seeds from a tomato plant. No amount of environmental input will yield an apple tree from those seeds, yet the “environment” (the amount of water, sun, pests, etc.) will affect which of the trajectories possible for a tomato plant – wilting, flourishing, or any of the stages in between – will be realized. Yet still the fact is that the *type* tomato is the only possible outcome, even though the particular token of a tomato we get may vary quite a bit.

Whether type-facts *really are* causally primary is a question to which we return in the next section, but the central claim of the current proposal – one that should stand notwithstanding possible points of disagreement within Nietzsche interpretation¹¹ – is that “type-facts,” which are present at birth, “play a powerful (but not exclusive) role in determining one’s behavior and values, though a far more powerful role than education or upbringing or conscious choice” (Knobe and Leiter 2007, 90).

This kind of view is largely unexplored in philosophical work in moral psychology. While upbringing and conscious choice are emphasized as contributing factors for moral motivation, the view that genetic propensities might be an essential component to, say the display of sincerity, is not countenanced. Yet there is overwhelming evidence from behavioral genetics in support of the thesis that personality traits are highly heritable. For example, according to a review of five studies conducted across five countries – with a sample size of 24,000 twins – an astonishing 60% of the variance in extraversion and 50% of the variance in neuroticism is estimated to be explained by genetic factors (astounding given that the average effect size (13–14%) identified in foundational studies in social psychological experiments (Loehlin 1992; cf. Knobe and Leiter, 92–3)).

To say that genetic propensities are an important (and neglected) factor in moral motivation is not to say that genes are fully determinative of behavior. As Knobe and Leiter (2007, 93) clarify,

[w]hen we say that a trait is heritable, we do not mean that it is produced entirely by a person’s genes, without any intervention from the environment. All we mean is that the differences between different people’s scores on this trait can be explained in part by differences in those people’s genetic material. This effect may not be direct. Differences in people’s genes might lead to differences in their environments, which in turn lead to differences in their scores on certain traits. Often the result will be a self-reinforcing cycle in which early behaviors that express a given trait lead the person to possess that trait to ever greater degrees. For example, a person’s initial extraverted behavior might leave her with a reputation for extraversion, which in turn makes her even more extraverted.

Accordingly, some caution is needed when dealing with Nietzsche’s talk of the “inevitability” or “necessity” of actions that follows from certain physiological or unconscious determinants:

It is utterly impossible that a person might *fail* to have the qualities and propensities of his elders and ancestors in his body: however much appearances might speak against it. This is the problem of race. If you know anything about the ancestors, you can draw conclusions about the child. Some sort of harmful immoderation, some sort of corner jealousy, a clumsy insistence on always being right – together, these three elements have constituted the true “vulgar” type in every age. And something like this will be passed on to the child just as certainly as contaminated blood. (BGE 264)

It should not be overlooked that Nietzsche is here referring to the heritability of *qualities* and *propensities*. Token identical propensities can generate different behavior depending on facts about the environment. So, when we say that traits are heritable, we mean “*broad* traits,” like extraversion, and neuroticism – “traits that produce a wide variety of different types of behavior” (Knobe and Leiter 2007, 95n) and the existence of which is nearly universally accepted among psychologists.

There is also relatively strong evidence of the heritability of aggressive anti-social behavior in children. For example, a heritability of 70% was found in a study of 1523 pairs of twins (Eley, Lichtenstein, and Stevenson 1999).¹² To say that a large causal role is played by genetic factors is not to say that there is a relationship of *causal necessity* between the possession of a trait and the display of certain behavior, but nor is it, *pace* Alfano (forthcoming), to say that it

is “normatively necessary” that certain behavior be displayed. Why not speak simply of tendencies? – tendencies *grounded* in genetics but nonetheless susceptible to influence from environmental factors?

The genetic etiology of the anti-social tendencies involved in psychopathology has also received recent attention. In a study of 626 pairs of 17-year-old male and female twins, significant genetic influence was found for two separate psychopathic traits: fearless dominance (66% phenotypic covariance) and impulsive anti-sociality (76% phenotypic covariance) (Blonigen *et al.* 2005). Early childhood education and the avoidance of certain experiences (e.g., bullying) are of considerable importance in managing the genetic influence of (especially a confluence of) psychopathic traits. In a fascinating interview with neuroscientist and pro-social psychopath James Fallon (Ohikuare 2014), Fallon offers a description of the insincerity with which he displays kind behavior, a description that nicely fits Nietzsche’s (BGE 264) (admittedly simplifying) claim that “‘education’ and ‘culture’ essentially *have* to be the art of deception – to deceive about lineage, about the inherited vulgarity in body and soul.” Discussing the positivity with which family members respond to his indirectly motivated but nonetheless beneficent behavior, Fallon says: “[y]ou’ve got to be kidding me. You accept this? It’s phony!”

7.3 Epiphenomenalism

The path is short from the acceptance of the Doctrine of Types to the acceptance of epiphenomenalism, as Leiter, and more recently, Riccardi argue. Let us start with Nietzsche’s phenomenological account of willing, which serves as independent motivation for the view that Nietzsche denies the causal efficacy of conscious acts of willing.¹³

In opposition to the popular view that the will is a unified thing sufficient for causing action, Nietzsche argues that phenomenological scrutiny reveals each act of willing to contain three components: (i) a commandeering thought; (ii) the feeling of bodily movement; and (iii) the meta-affect of power (BGE 19). The meta-affect of power is generated by the agent’s *identification* with the commandeering thought, which, owing to its temporal priority to the bodily movement, *seems* sufficient for causing the bodily movement. I take myself to be a free and efficacious being in identifying with the thought, for example, “I will push the button.” The “affect of superiority” that we feel in identifying with the thought-component of acts of willing is misguided because *thoughts themselves* are not preceded by the experience of willing. As Nietzsche observes “a thought comes when ‘it’ wants, and not when ‘I’ want,” (BGE 17) which is to say, the phenomenological basis for thinking “I” the thinker am the cause of my actions, cannot extend to the very mental acts that are supposed to do the causal work. In short, the phenomenology of willing is no guide to the causal explanation of either thought or action.

One take-away point that Leiter (2009, 113) extracts from the above argument is “that actions that follow upon our experience of willing (which includes those thoughts) are not caused in a way sufficient to underwrite ascriptions of moral responsibility.” *But*, we need not think that willing must ground moral responsibility in order for it to ground *causal responsibility*. Nietzsche’s embrace of hard incompatibilism – commitment to which we cannot outline here – need not rule out taking conscious acts of willing to be causally efficacious of action. As hard incompatibilist Pereboom (1995, 31) puts it,

[soft determinists like] Ayer and Dennett, among others, have pointed out that the determination of our deliberations, choices, actions, and their consequences does not undermine their causal efficacy. The hard determinist can legitimately appropriate this position. It is true that according to hard determinism we are not free in the sense required for moral responsibility, and therefore, what

happens cannot be affected by choices that are free in this sense. But what happens may nevertheless be caused by the deliberations we engage in and the choices we make.

Since moral responsibility is unnecessary for causal responsibility, let's turn to a more direct argument for epiphenomenalism from the Doctrine of Types.

Nietzsche thinks we mistakenly posit a causal relationship between two correlated events – for example, eating a meager diet, and being slender – neglecting that *both* are caused by a single factor lying beyond one's immediate control, for example, a fast metabolism (TI VI: 1). The latter is the *deep cause* and its hiddenness from introspection misleads us into thinking that one of its effects (eating a meager diet) is the freely chosen cause of the deep cause's *other* effect (being slender):

Everybody knows the book of the famous Cornaro in which he recommends his slender diet as a recipe for a long and happy life...The worthy Italian thought his diet was the *cause* of his long life, whereas the precondition for a long life, the extraordinary slowness of his metabolism, the consumption of so little, was the cause of his slender diet. He was not free to eat little *or* much; his frugality was not a matter of "free will": he became sick when he ate more. (TI VI: 1)

Morality and religion, on Nietzsche's view, are guilty of the same mistake expressed in "Cornaro's error" – they recommend certain practices as the causes of a happy life, ignoring the fact that one's susceptibility to both religious practice and its pleasures is itself the product of certain type-facts:

[P]eople experience hope *because* their fundamental feeling is strong and rich again; people have faith in God *because* the feeling of strength and peace gives them peace. – Morality and religion can be exhaustively accounted for by *the psychology of error*: in every single case, cause and effect are confused; or truth is confused with the effects of *believing* that something is true; or a state of consciousness is confused with its causes. (TI VI: 6)

Consciously followed prescriptions cannot be causally efficacious of action if such actions are to be explained by one's unconscious psychological and physiological constitution. While commitment to epiphenomenalism was once counted as a *reductio* for one's view of the mind-body relation, there is considerable empirical evidence suggesting that the "conscious will" may in fact be an illusion. Following Libet, whose studies revealed that the brain's electrical activity (or "readiness potential") is temporally prior to the subject's conscious experience of willing, Wegner's work suggests that both actions and experiences of willing are products of unconscious causes. On Wegner's model, there is *no* causal connection between the experience of willing and action (Wegner 2002, 68, 98). More recent work by Desmurget *et al.* (2009, 811–13) suggests that parietal lobe stimulation underlies the experience of conscious willing, and moreover, that the experience of willing is neither necessary nor sufficient for actual bodily movement. This psychological literature appears to support Nietzsche's (TI VI: 3) counterintuitive view that "[t]he 'inner world' is full of illusions and phantasms: will is one of them. The will does not do anything anymore, and so it does not explain anything anymore either – it just accompanies the process, but it can be absent as well."

Caution is in order here, however. More empirical work is required before we can rule out that the brain manipulation involved in the above experiments leaves intact the ordinary causal influences of intention. It is possible that brain stimulation generates feedback of success in a way that undercuts the *ordinary* causal chain between experiences of willing, feedback of success, and resultant bodily movement.¹⁴

Apart from the empirical underdetermination of epiphenomenalism, there may be grounds for doubting that Nietzsche holds the strong thesis that conscious mental states are *never* causally efficacious. As Katsafanas (2005, 21) argues, Nietzsche claims that when bad conscience (an unconscious attitude) gets consciously conceptualized as *guilt*, this conceptualization has a causal impact on the nature of the individual's drives, that is, her unconscious mental states. On Katsafanas' view, it is in virtue of conceptualization as such that conscious states can have a causal impact on the "mental economy" of the whole individual, including her drives. Riccardi (forthcoming) convincingly argues that Nietzsche thinks a low-level unconscious conceptualization "governs our perception," and given that this "imagistic" conceptualization is unconscious, we cannot accept as stated Katsafanas' claim that conscious mental states are causally efficacious *because* they conceptualize previously nonconceptual content.

According to both Leiter and Riccardi, the causal impotence of conscious states follows from the "explanatory priority" of type-facts: *no genuine explanatory power* can be given to other facts about the person. As Leiter (2001, 299) puts it, "[i]f type-facts determine a person's 'ideas and values' then even if 'ideas and values' determine one's actions, these actions and choices themselves are all the necessary consequences of the underlying type-facts." Notice that Leiter (2001, 294) admits that "a person's conscious states may be part of the causal chain leading up to action," but that "they play that role only in virtue of type-facts about the person." That is, Leiter thinks consciousness is *kind-epiphenomenal*. Riccardi (forthcoming) similarly claims that "the real psychological causality is at the level of unconscious states Nietzsche conceives of as drives."

Should we, however, follow Leiter and Riccardi in claiming that consciousness is kind-epiphenomenal? This question is pressing given that Riccardi (forthcoming) reasonably concedes Katsafanas' point that self-conscious (or linguistic) representations often have an "enormous impact" on one's drives, *whilst nonetheless denying* that this impact is *causal*. Riccardi claims that conscious states, including beliefs involved in *linguistically* conceptualizing bad conscience as guilt, are nonetheless kind-epiphenomenal since, "[o]nly once this belief has been internalized would one's state be effectively conceptualized as 'guilt' and become behaviorally relevant as such. The crucial factor, thus, is the kind of psychological mechanism responsible for the internalization." The psychological mechanism responsible for internalization is surely part of the sub-personal story. Yet, this fact by itself is insufficient for adjudicating between Epiphenomenalists (Leiter and Riccardi) and *Conscious Efficacy Theorists* (like Katsafanas). That is, the Conscious Efficacy Theorist will concede that in order to have an impact on one's drives, conscious states need to be integrated into the agent's subpersonal mechanisms – after all, a drive is a kind of subpersonal disposition – but why should this entail that the *real* nexus of causal relations is found on the sub-personal level? It might seem to be, of course, if the sub-personal level were *always* sufficient for the effect, even in the absence of the conscious state. This will, in turn, raise some difficult issues about causation.

Absent an account of Nietzsche's views on causation – which we have reason to believe are sometimes rather exotic¹⁵ – and argument against the scientifically responsible incompatibilist who is also a *Conscious Efficacy Theorist* (e.g., Pereboom), let us attribute to Nietzsche a thesis weaker and more provisional than Epiphenomenalism, but which nonetheless shares part of Epiphenomenalism's motivation: While conscious states can be efficacious of action, frequently the real causes of action are hidden from introspection, and are often retrospectively confabulated. While we cannot defend the following claim here, it is plausible that Nietzsche takes the efficacy of one's conscious states to be proportionate to her self-control. We *can*, however, provide an account of Nietzsche's underappreciated claims about the nature of self-control.

7.4 Nietzsche's Strength Model of Self-Control

Nietzsche's denial that our wills are free is closely connected to his denial of moral responsibility:

Now one finally discovers that [human] nature, too, cannot be accountable, inasmuch as it is altogether a necessary consequence and assembled from the elements and influence of things past and present: that is to say, that man can be made accountable for nothing, not for his nature, nor for his motives, nor for his actions, nor for the effects he produces. One has thereby attained to the knowledge that the history of the moral sensations is the history of an error, the error of accountability, which rests on the error of freedom of will. (HAH 39)

Similar passages can be found throughout Nietzsche's corpus (GM Pref: 2; BGE 19; BGE 21; TI VI: 3; TI VI: 7; D 148) and empirical support has recently been levied in favor of Nietzsche's concomitant view that the pervasiveness of belief in free will is to be explained in large part by our punitive motivations (Clark *et al.* 2014).¹⁶ Yet, in spite of his consistently hard incompatibilist stance on free will, Nietzsche thinks there is another question we can ask about a person's will, namely one about its *strength*.

That Nietzsche takes this question to be of practical significance is clear from his claim that "in real life," the question of the status of one's will "is only a matter of *strong* and *weak* wills. It is almost a symptom of what is lacking in a thinker *when he senses some compulsion, need, having-to-follow, pressure, unfreedom* in every 'causal connection' and 'psychological necessity'" (BGE 21, emphasis added).

The denial of freedom and responsibility, in other words, does not rule out differences between "weak" wills and those that display (or are disposed to display) "self-control." Talk of "self-control" need not imply that the self is anything like a *self-cause*, or even capable of reflective-detachment from one's drives (cf. Katsafanas 2013, 750–2). On the contrary, Nietzsche holds that type-facts function as enabling conditions for the display of strength of will. For example, In *Twilight*, Nietzsche claims that to "stay true to *my type*" the ability to resist impulses is indispensable. This disposition, Nietzsche continues,

is the *first* preliminary schooling for spirituality: *not* to react immediately to a stimulus, but instead to take control of inhibiting, excluding instincts. Learning to *see*, as I understand it, is close to what an unphilosophical way of speaking calls a strong will: the essential thing here is precisely *not* "to will," to be *able* to suspend the decision. Every characteristic absence of spirituality, every piece of common vulgarity, is due to an inability to resist a stimulus – you *have* to react, you follow every impulse. In many cases this sort of compulsion is already a pathology... (TI VII: 6)

Nietzsche here identifies "strength of will" with that in virtue of which individuals like he – "affirmative" types, he adds – can resist stimuli and remain "true to their type." Since we know that Nietzsche counts himself among "legislators of value," that is, philosophers in the honorific sense (BGE 211), it should not surprise us that he puts a high premium on self-control. His distaste for the *laissez-aller* is due to the conviction that commitment to long-term goals – "obedience for a long time in a single direction" – is a necessary (though insufficient) condition for the production of "great things," those that make life worth living (BGE 188). It is for this reason that Nietzsche (BGE 212) claims that, "strength of will and the hardness and capacity for long-term resolution must belong to the concept of 'greatness,' in the philosopher's ideal."

Nietzsche's strength model of self-control has recently found support in a growing body of empirical research (Baumeister *et al.* 1998; Baumeister and Vohs 2007; Duckworth 2011;

Muraven, Tice, and Baumeister 1998). The evidence suggests that self-control is best understood as neither a skill nor kind of knowledge, but as the exertion of a depletable supply of energy. Several studies suggest that impulse control (e.g., resisting the temptation to eat a desired candy) markedly decreases one's likelihood to persevere in subsequent tasks of self-control (e.g., puzzle-solving problems). The evidence further suggests that we exercise a *common resource* in delaying gratification, monitoring impulses, and other controlled processes (Baumeister *et al.* 1998, 1253).

The “strength of energy model” is not to be construed as an essentially noncognitive or brute force. It involves mental effort, and as Holton (2009, 212) claims, functions to resist reconsiderations of our resolutions;¹⁷ strength of will is thus responsive to reasons. As Neil Levy (2010, 271, drawing on Baumeister *et al.* 2008) notes, “when subjects are reminded of their values, or offered cash incentives, they are able to hold out for longer periods against the urge to succumb.” This is not, Nietzsche will claim, because the “self” of self-control is an essentially rational or unified thing (BGE 19). Rather, given that Nietzsche takes there to be a knowledge drive, we need only accept the plausible idea that this drive will be co-opted in many effective exercises of self-regulation. To deny this is to deny that drive psychology is compatible with instrumental rationality and Nietzsche does no such thing.

Having accepted that self-control is partly cognitive, it is tempting to suppose that one is *responsible* for at least some feats of resolution. But this temptation is apt to misguide. As Nietzsche claims after introducing six techniques for the “self-mastery” of overzealous drives (D 109), “while ‘we’ believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about another.*” Talk of self-control, like Nietzsche’s talk of self-overcoming (D 192; BGE 61; BGE 257; TI X: 38), strongly suggests the ability for responsibility-accruing acts of reflective-detachment, but Nietzsche’s recurring insistence on the fundamentality of drive psychology invites us to look for naturalistic explanations for wonderful feats of achievement, aesthetic and otherwise. Fortunately, for Nietzsche and Nietzscheans, the beginnings of such explanations are in the offing.¹⁸

7.5 Conclusion

What can explain Nietzsche’s seeming prescience about moral psychological truths? While lacking many of the methodological resources for *confirming* the truth of the views he endorses, Nietzsche was nonetheless privy to a variety of resources that would support plausible speculative hypotheses.¹⁹ One such source is keen introspective observation: recall for instance Nietzsche’s (BGE 19) intricate phenomenological discussion of the experience of willing. In addition to carefully examining his own mental life, Nietzsche was a voracious reader of our keenest observers of human motivation: the realist history of Thucydides, the aphorisms of La Rochefoucauld, the prose of Stendhal – whom Nietzsche honors as “France’s last great psychologist” (BGE 254) – all contribute to Nietzsche’s insight into what moves us. Nietzsche’s effort to provide naturalistic accounts of human behavior owes much to the influence of the use of physiological explanation by German Materialists, and his interest in questions of agency and evaluative attitudes bears witness to his indebtedness to Schopenhauer (Leiter 2015, 42–50).

Nietzsche’s 1886 claim that “[a]ll psychology so far has been stuck in moral prejudices and fears” (BGE 23) remains partly true of the vast majority of philosophical literature on moral psychology. Antiquarian prejudices continue to attract philosophers to the armchair. To these figures, we offer Nietzsche’s (EH II: 1) injunction to “[s]it as little as possible; do not believe any idea that was not conceived while moving around outside....” Today, Nietzsche’s claim that psychology is the *queen of the sciences* demands that we move toward attending to the results of *actual* psychology.²⁰

Notes

- 1 We start with translations by Kaufmann, Hollingdale, Norman, and/or Clark & Swensen, making modifications based on Friedrich Nietzsche, *Sämtliche Werke: Kritische Studienausgabe in 15 Bänden*, edited by G. Colli and M. Montinari (Berlin: de Gruyter, 1980). Roman numerals refer to major parts or chapters in Nietzsche's works; Arabic numerals refer to sections not pages. We 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 Genealogy of Morals* (GM); *Human, All Too Human* (HAH); *Thus Spoke Zarathustra* (Z); *Twilight of the Idols* (TI); *The Will to Power* (WP).
- 2 Earlier in same text, Nietzsche claims (BGE 6) that the philosopher's "morals 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."
- 3 In this context Nietzsche's talk of morality as symptomatic of *physiological* processes amounts to the same commitment, namely to the explanation of normative judgments in terms of nonrational (and often, sub-personal) natural phenomena. This form of explanation in Nietzsche marks the influence of the German Materialists (Leiter 2015, 50–6) and also of the role of affective force in Spinoza's mind-body parallelism, familiarity with which Nietzsche gained through Kuno Fischer's work; see, for example, Wollenberg (2013, 625–44).
- 4 For a useful overview of various sentimentalist positions, see chapter 1 of Prinz (2007). The thesis we attribute to Nietzsche is in the first place one about the *genesis* of moral judgments, and so, talk of 'Nietzsche's sentimentalism' must be sharply separated from normative sentimentalist theories, for example, Slote (2010).
- 5 Nietzsche's (*rather promiscuous*) ontology of drives should be approached with due caution – see Leiter (2013, 249 fn.) for discussion.
- 6 Recent evidence that we perceive ordinary objects as *micro-valenced* might prove useful in understanding how our affective dispositions can heighten the valence with which we perceive objects. See Lebrecht *et al.* (2012, 107).
- 7 Kant *does* accord a central moral function to the emotion of "respect," but since respect has as its object "the moral law," Kant's moral psychology is far too intellectualist to earn support from growing evidence in favor of noncognitivism. As Bagnoli (2014) compellingly argues, Kant cannot be charged with ignoring moral phenomenology, but Nietzsche would nonetheless object to the purportedly ahistorical status of our phenomenology of moral law.
- 8 Affective defect sheds light on more than the nature of our *moral* responses. The nature of practical reasoning is also elucidated by research on certain affective abnormalities (Damasio 1994). As Arpaly (2003, 59) helpfully reads an upshot of Damasio's research, "[b]rain-damaged nonfeelers, despite an unharmed ability to deliberate and reflect, make bad decisions because they are denied [ordinary] feeling-based access to their own background knowledge in making those decisions." Nietzsche's drive psychology finds support in the view that ordinarily, one's deliberative frame is bounded and guided by emotional cues (*perhaps* due to a mechanism similar to that operative with perceptual micro-valences – see fn. 6).
- 9 For criticism of Haidt's overreaching, see Leiter (2013, 256). The mechanism for this post-hoc "justification" might be provided by research on cognitive dissonance, according to which tension between our beliefs and actions impels us to adjust our attitudes in such a way that ameliorates inconsistency and places us in favorable light. For the original presentation of the theory, see Festinger (1957) and for a comprehensive treatment, see Cooper (2007).
- 10 Thanks to an anonymous reviewer for seeking clarification on this point.
- 11 Cf. Gemes and Janaway (2005) and Leiter (2015, 256–7).
- 12 The results of other studies on the genetic basis of violent behavior in children were lower but still too high to be written off as either experimental artifacts or measurement errors: 60%, Edelbrock *et al.* 1995; 49%, Deater-Deckard and Plomin 1999; 60%, Schmitz, Fulker, and Mrazek 1995.
- 13 By "conscious" we mean neither phenomenal consciousness nor awareness, but "self-conscious," since as Riccardi (forthcoming) notes, Nietzsche describes one's 'becoming conscious' in terms of "seeing itself in the mirror" (GS 354), and his sense of "consciousness" is unique to humans, while phenomenal consciousness and awareness are not.

- 14 Thanks to Gideon Yaffe for bringing this challenge to our attention.
- 15 For example, Nietzsche's (WP 551) claim that "*causa efficiens* and *causa finalis* are fundamentally one". For helpful discussion, see Poellner (1995, 30–46).
- 16 The apt passage with which Clark *et al.* open their paper continues thus: "The notion of will was essentially designed with punishment in mind, which is to say the desire to *assign guilt*. The whole of ancient psychology, the psychology of will, was conditioned by the desire of its architects (the priests at the head of the ancient community) to establish their *right* to inflict punishment – or to assign the right to God...People were considered 'free' so that they could be judged and punished – " (TI VI: 7). Nietzsche's claim may even be historically accurate: see Frede (2011, 115).
- 17 It is an interesting question how to understand the relationship between strength of will and desire. On Holton's (2009) view, the resources of will power are drawn on to resist desire and promote a resolution, which (being a kind of intention) is a motivational state, yet irreducible to a combination of desire and belief.
- 18 Were Nietzsche to collaborate with social psychologists engaged in self-control research, it is likely that he would further pursue studies of the sort that underlie Duckworth's (2011, 2639) claim that "[a]lthough older individuals are, on average, more self-controlled than younger individuals there are nonetheless salient differences in self-control among individuals of the same age." One way to hear Nietzsche's frequent "invocations" to self-overcoming is as causal contributors to the optimization (*though not necessarily*, maximization) of self-control and related dispositions that persons of certain type-facts harbor.
- 19 See Leiter (2015, 263).
- 20 We thank the editors and an anonymous referee for helpful comments on an earlier version of this chapter.

References

- Alfano, Mark. Forthcoming. "An enchanting abundance of types: Nietzsche's modest unity of virtue thesis." *Journal of Value Inquiry*.
- Arpaly, Nomy. 2003. *Unprincipled Virtue: an Inquiry into Moral Agency*. New York: Oxford University Press.
- Bagnoli, Carla. 2014. "Moral Objectivity: A Kantian Illusion?" *Journal of Value Inquiry*. DOI: 10.1007/s10790-014-9448-7.
- Baumeister, R.F., E. Bratslavsky, M. Muraven, and D.M. Tice. 1998. "Ego Depletion: Is the Active Self a Limited Resource?" *Journal of Personality and Social Psychology*, 74: 1252–1265. DOI: 10.1037/0022-3514.74.5.1252.
- Baumeister, R.F., E.A. Sparks, T.F. Stillman, and K.D. Vohs. 2008. "Free Will in Consumer behavior: Self-control, ego depletion, and choice." *Journal of Consumer Psychology*, 18: 4–13. DOI: 10.1016/j.jcps.2007.10.002.
- Baumeister, R.F. and K.D. Vohs. 2007. "Self-Regulation, Ego Depletion, and Motivation." *Social and Personality Psychology Compass*, 1: 115–128. DOI: 10.1111/j.1751-9004.2007.00001.x.
- Blair, James. 1995. "A cognitive developmental approach to morality: Investigating the psychopath." *Cognition*, 57: 1–29. DOI: 10.1016/0010-0277(95)00676-P.
- Blonigen, D. M., B.M. Hicks, R.F. Krueger, C.J. Patrick, and W.G. Iacono. 2005. "Psychopathic personality traits: heritability and genetic overlap with internalizing and externalizing psychopathology." *Psychological Medicine*, 35: 637–648. DOI: 10.1017/S0033291704004180.
- Clark, C.J., J. Luguri, P.H. Ditto, J. Knobe, A.F. Shariff, and R.F. Baumeister. 2014. "Free to Punish: A Motivated Account of Free Will Belief." *Journal of Personality and Social Psychology*, 106 (4): 501–513. DOI: 10.1037/a0035880.
- Cooper, Joel. 2007. *Cognitive Dissonance: 50 Years of A Classic Theory*. London: Sage.
- Cushman, E., L. Young, and J.D. Greene. 2010. "Multi-System Moral Psychology." In *The Moral Psychology Handbook*, edited by J. Doris and the Moral Psychology Research Group, 47–71. Oxford: Oxford University Press.
- Damasio, Antonio. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam's Sons.

- Deater-Deckard, K. and R. Plomin. 1999. "An Adoption Study of the Etiology of Teacher Reports of Externalizing Problems in Middle Childhood." *Child Development*, 70: 144–154. DOI: 10.1111/1467-8624.00011.
- Desmurget, M., K.T. Reilly, N. Richard, A. Szathmari, C. Mottolose, and A. Sirigu. 2009. "Movement Intention After Parietal Cortex Stimulation in Humans." *Science*, 324 (5928): 811–813. DOI: 10.1126/science.1169896.
- Duckworth, Angela. 2011. "The Significance of Self-control." *Proceedings of the National Academy of Sciences of the United States*, 108(7): 2639–2640. DOI: 10.1073/pnas.1019725108.
- Edelbrock, C., R. Rende, R. Plomin, and L.A. Thompson. 1995. "A Twin Study of Competence and Problem Behavior in Childhood and Early Adolescence." *Journal of Child Psychology and Psychiatry*, 36(5): 775–785. DOI: 10.1111/j.1469-7610.1995.tb01328.x.
- Eley, T., P. Lichtenstein, and J. Stevenson. 1999. "Sex Differences in the Etiology of Aggressive and Nonaggressive Antisocial Behavior: Results from Two Twin Studies." *Child Development*, 70: 155–168. DOI: 10.1111/1467-8624.00012.
- Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Frede, Michael. 2011. *A Free Will: Origins of the Notion in Ancient Thought*, edited by A.A. Long. Berkeley: University of California Press.
- Gemes, Ken and Christopher Janaway. 2005. "Naturalism and Value in Nietzsche." *Philosophy and Phenomenological Research*, 71(3): 729–740. DOI: 10.1111/j.1933-1592.2005.tb00485.x.
- Goldie, Peter. 2002. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.
- Greene, Joshua. 2007. "The Secret Joke of Kant's Soul." In *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Disease, and Development*, edited by W. Sinnott-Armstrong, 35–117. Cambridge, MA: MIT Press.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Vintage Books.
- Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review*, 108: 814–834. DOI: 10.1037//0033-295X.108.4.814.
- Hoffman, Martin. 2000. *Empathy and Moral Development: Implications for Caring and Justice*. New York: Cambridge University Press.
- Holton, Richard. 2009. *Willing, Wanting, Waiting*. New York: Oxford University Press.
- Katsafanas, Paul. 2013. "Nietzsche's Philosophical Psychology." In *The Oxford Handbook of Nietzsche*, edited by J. Richardson and K. Gemes, 727–754. New York: Oxford University Press.
- Katsafanas, Paul. 2005. "Nietzsche's Theory of Mind: Consciousness and Conceptualization." *European Journal of Philosophy*, 13: 1–31. DOI: 10.1111/j.0966-8373.2005.00220.x.
- Knobe, Joshua and Brian Leiter. 2007. "The Case for Nietzschean Moral Psychology." In *Nietzsche and Morality*, edited by B. Leiter and N. Sinhababu, 83–109. New York: Oxford University Press.
- Lebrecht, S., M. Bar, L.F. Barrett, and M.J. Tarr. 2012. "Micro-valences: perceiving affective valence in everyday objects." *Frontiers in Psychology*, 3(107): 1–5. DOI: 10.3389/fpsyg.2012.00107.
- Leiter, Brian. 2013. "Moralities are a Sign-Language of the Affects." *Social Philosophy & Policy*, 30: 237–258. DOI: 10.1017/S0265052513000113.
- Leiter, Brian. 2015. *Nietzsche on Morality*, 2nd ed. London: Routledge.
- Leiter, Brian. 2009. "Nietzsche's Theory of the Will." In *Nietzsche on Freedom and Autonomy*, edited by K. Gemes and S. May, 107–126. Oxford University Press.
- Leiter, Brian. 2001. "The Paradox of Fatalism and Self-Creation in Nietzsche." In *Nietzsche: Oxford Readings in Philosophy*, edited by B. Leiter and J. Richardson, 281–321. New York: Oxford University Press.
- Levy, Neil. 2010. "Addiction and Compulsion." In *A Companion to the Philosophy of Action*, edited by T. O'Connor and C. Sandis, 267–273. Oxford: Wiley-Blackwell.
- Loehlin, John. 1992. *Genes and Environment in Personality and Development*. Newberry Park, CA: Sage.
- Muraven, M., D.M. Tice, and R.F. Baumeister. 1998. "Self-control as limited resource: Regulatory depletion patterns." *Journal of Personality and Social Psychology*, 74: 774–789. DOI: 10.1037/0022-3514.74.3.774.
- Nichols, Shaun. 2004. *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York: Oxford University Press.
- Ohikuare, Judith. 2014. "Life as a Nonviolent Psychopath." *The Atlantic*, January 21. Accessed August 18, 2014. <http://www.theatlantic.com/health/archive/2014/01/life-as-a-nonviolent-psychopath/282271/>

- Pereboom, Derk. 1995. "Determinism *al Dente*." *Noûs*, 29(1): 21–45.
- Pereboom, Derk. 2001. *Living Without Free Will*. New York: Cambridge University Press.
- Poellner, Peter. 1995. *Nietzsche and Metaphysics*. Oxford: Oxford University Press.
- Prinz, Jesse. 2007. *The Emotional Construction of Morals*. New York: Oxford University Press.
- Reginster, Bernard. 2006. *The Affirmation of Life: Nietzsche on Overcoming Nihilism*. Cambridge, MA: Harvard University Press.
- Riccardi, Mattia. Forthcoming. "Nietzsche on the Superficiality of Consciousness." In *Nietzsche on Consciousness and the Embodied Mind*, edited by M. Dries. Berlin/Boston: De Gruyter Press.
- Schmitz, S., D.W. Fulker, and D.A. Mrazek. 1995. "Problem Behavior in Early and Middle Childhood: An Initial Behavior Genetic Analysis." *Journal of Child Psychology and Psychiatry*, 36: 1443–1458. DOI: 10.1111/j.1469-7610.1995.tb01674.x.
- Sommers, Tammler. 2007. "The Objective Attitude." *The Philosophical Quarterly*, 57(228): 321–341. DOI: 10.1111/j.1467-9213.2007.487.x.
- Slote, Michael. 2010. *Moral Sentimentalism*. New York: Oxford University Press.
- Wegner, Daniel. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wollenberg, David. 2013. "Nietzsche, Spinoza, and the Moral Affects." *Journal of the History of Philosophy*, 51(4): 617–649. DOI: 10.1353/hph.2013.0076.