Original Article

Can Character Traits Be Based on Brute Psychological Facts?

Iskra Fileva
Department of Philosophy, University of Colorado at Boulder, Boulder, CO 80309-0232

Correspondence
Iskra Fileva, Department of Philosophy, University of Colorado at Boulder, Boulder, CO 80309-0232, U.S.A.
Email: iskra.fileva@gmail.com

Abstract
Some of our largely unchosen first-order reactions, such as disgust, can underwrite morally-laden character traits. This observation is in tension with the plausible idea that virtues and vices are based on reasons. I propose a way to resolve the tension.

1. INTRODUCTION

A person's character traits – that is, her virtues and vices[1] – seem deeply intertwined with her practical reasoning. Consider two arguments in support of this claim. First, if a person's actions are not motivated by reasons at all, character-based explanations of those actions become inappropriate. Thus, a phobic person with an incapacitating fear of heights is not considered cowardly when she refuses to get on a plane, while a schizophrenic with hallucination who attacks another in the belief the other is about to kill him is not said to show the vices of hostility and aggression.[2]

Second, in attributing any particular trait to a person, we seem to ascribe to her sensitivity to particular kinds of reasons. Thus, the kind person is disposed to take the feelings and well-being of others as reasons for action, while the selfish one is attuned to reasons that have to do with her own interests. Similarly, the courageous person takes risks that are worth taking – i.e., such as there is a reason to take. A person who takes risks he or she has no *reason* to take – e.g., drives without a seatbelt – is foolhardy, rather than courageous. The virtue of courage, thus, requires responsiveness to reasons to act bravely.

Aristotle makes a similar point both with regard to bravery and with regard to the connection between character and reason. He has this to say about a person's propensity not to feel fear and that propensity's connection to bravery:

> [F]or to fear some things is even right and noble, and it is base not to fear them – e.g. disgrace; he who fears this is good and modest, and he who does not is shameless.[3]

And he goes on to define character excellence thus:

---

[1] It could be suggested that there are morally neutral character traits, such as shyness. However, I prefer to reserve the label 'character trait' for morally-laden traits, and to refer to morally neutral traits as 'personality traits.'
[2] See also Fileva (2016).
[3] Aristotle (2003/349 B.C., 1115a11–15; my italics).

> Excellence [of character], then, is a state concerned with choice, lying in a mean relative to us, *this being determined by reason and in the way in which the man of practical wisdom would determine it*.[4]

Plato's account of character, at least in the *Republic*, is similar to Aristotle's: in Book IV of the *Republic*, Plato argues that virtue is a matter of governance of the non-rational part of the soul by the rational part.[5] Kant aligned character with reason very closely: for Kant, virtue is a matter of will to fulfill one's duties.[6] We find the idea of virtue as perception of reasons in a number of contemporary authors as well.[7]

Traits, thus, appear to differ from attitudes and dispositions based entirely on brute psychological facts, such as aerophobia, or a propensity to avoid gas stations and buses because one is averse to the smell of gasoline. For in contrast with the former, the latter are largely *disconnected* from reasons and reasoning. It is precisely because we take character traits to be, in some very clear sense, an expression of a person's reasoning and we do not take purely psychological conditions such as phobias and aversions to be such an expression that we hold people responsible for their traits, but not for their purely psychological propensities. Thus we say, 'Don't be such a pushover!' or, 'Stop being such a coward,' but not, 'Enough with the aversion to the smell of gasoline!' Or, 'Get rid of that phobia! Get over yourself!'[8]

Relatedly, we suppose traits to be responsive to reasons in ways in which attitudes based on brute psychological facts are not. For instance, if a person is unkind, I may reason with her, attempt to convince her that kindness is important, and I may succeed. Even if I fail, engaging in an argument with regard to the matter seems appropriate. It would certainly be odd for the other to say, 'Well, this is not really a matter of reasons, I just find kindness nauseating. I dislike it in much the way you dislike the smell of gasoline.' If, by contrast, someone likes the smell of gasoline, it would appear misguided of me to try to *convince* him that it is a smell one ought to be averse to.

So far, so good. The problem is that some traits seem based on brute psychological facts, such as first-order attractions and aversions. Thus, in Amelie Nothomb's *Hygiene and the Assassin*, reclusive writer Prétextat Tach is portrayed as being extremely sexist on purely psychological grounds: Tach's sexism stems from an aversion to the characteristically female attributes of an adult woman's physique. Tach has this to say to a female journalist interviewing him about the laughter of women:

---

[4] Ibid., 1106b36–1107a3.

[5] Some commentators have suggested that for Plato, virtue is *solely* a matter of knowledge (the 'virtue is knowledge' doctrine) while vice is *solely* a matter of ignorance. Paul Gooch (1971, 124) ascribes this interpretation to, for example, Eric Dodds (1945) and Michael O'Brien (1976). Gooch himself opposes this interpretation. I am inclined to oppose it as well: I think that Plato's account of virtue evolved over time – it is more intellectualist in the *Protagoras*, less so in the *Republic*, and perhaps even less so in the *Sophist*. But this exegetical point does not matter for my purposes – for Plato, virtue is always at least *partly* a matter of knowledge of what is right and good (and thus, has a necessary connection to reason), whether or not it is *exclusively* a matter of such knowledge.

[6] Kant writes, 'Virtue is, therefore, the moral strength of a *human being's will* in fulfilling his *duty*, a moral *constraint* through his own lawgiving reason' (1991/1797, 6:405).

[7] See, e.g., McDowell (1979) and Murdoch (2001).

[8] I should probably note here that Hume's view of character may be thought not to fit this general framework since Hume ties virtues of character to sentiment rather than reason. Hume (1975/1748) suggests that virtues are mental dispositions to have sentiments that move one to act, where the actions (and so, the qualities of mind which give rise to them) would occasion pleasure and approbation in an observer who adopts the 'common point of view.' If Hume's proposed route of understanding virtues and vices were to be followed, the question becomes *why* psychological conditions such as phobias do not give rise to moral approbation and disapprobation. Hume's general idea is that we approve of traits useful to either society or their possessor or both. This may be so, but it does not give us the answer to our question: phobias are typically useful neither to their possessor nor to society, but they are not, on that account, vices. The question is: why?

> Uncontrollable laughter. Yet another female ailment. I've never seen a man double over the way women do in these cases. It must come from the uterus. Everything disgusting in life comes from the uterus. Little girls do not have a uterus, I don't think, or if they do, it's a toy, a parody of a uterus. Little girls should be killed the moment their fake uterus becomes real, to spare them the type of terrific, painful hysteria you are suffering from at this very moment.[9]

So great is Tach's aversion to the defining properties of a woman's body, indeed, that in his early youth, he killed his prepubescent girlfriend on the day of her first period so she would not be 'defiled' by blood. He then let her body float in the river, to be cleansed by the water (hence, the book's title 'Hygiene and the Assassin').

Varieties of racism too may be based on such brute psychological facts. Thus, in Tarantino's 2009 film *Inglorious Basterds*, a Nazi officer explains to a French farmer why he and other Nazis hate Jews:

> When you see a squirrel running across your yard, do you feel you must kill it? No. But if a rat comes close to you or runs across your room, then what? Hate. You want to destroy it. Right? No real reason, we don't feel that way about the squirrel. In fact, they are basically the same animal less the nice tail. We just hate the rat. This explains the Jew. We don't know why we hate them, we just do.

Nor are character traits whose motivational component consists in psychological facts an exclusive prerogative of fictional characters: real-world homophobes, for instance, frequently claim that they find homosexual acts disgusting,[10] and homophobic murderers have even argued in court that their own disgust should be taken as a mitigating factor in the offences committed.[11] This gives rise to a question: *can* morally-laden traits can be based on brute psychological facts about their possessor, rather than on reasons? If they can, how if at all do traits differ from purely psychological attitudes and dispositions, such as phobias and revulsions? If they do not differ, why do we treat them differently? These are the questions I propose to answer here. I will argue that there is indeed a difference between character traits and the purely psychological aspects of someone's personality, and one which can help justify the differential treatment. But that difference, I will further suggest, is not a difference in motivation, although I will qualify this claim.

One brief clarification is due before we proceed further. Traits, as I have argued elsewhere, have both a motivational and a behavioral component.[12] A kind person is motivated to behave kindly, but she also does behave kindly, at least sometimes – if she is motivated to be kind but, for one reason or another, never actually shows any care for others (say, because she wants to be perceived as 'tough'), she cannot be said to be a kind person. Upon occasion, the behavioral component of a trait may be nil, and a trait may be ascribed to someone purely on the basis of motivation, as when we say that someone is secretly jealous of his spouse or secretly envious of his

---

[9] Nothomb (2010).

[10] Martha Nussbaum (2010) discusses in detail this tendency to take disgust as a reason for legal restrictions on the disgust-provoking individuals and behaviors – even when the restrictions would violate individual rights. Her view of disgust as a distorting influence on moral judgment has been disputed. See, for example, Plakias (2013); Deigh (2006); and Clark & Fessler (2015). For my purposes, however, this objection does not matter. It may well be that disgust has not only a useful physiological but a useful moral function, and that it helps attune us to certain sorts of considerations. Perhaps, the capacity to experience moral repugnance is rooted in the ability to feel disgust. The question that interests me is: what of dispositions based on disgust *disconnected* from reasons? On this point, see also footnote 30 below.

[11] See *Commonwealth v. Carr* (1990, 1363). One judge even accepted this argument and reduced the sentence of the defendant. See Belkin (1988, 8). Both cases are discussed in Kahan (1988, 1621–1624). The essay is a review of Miller (1997).

[12] See Fileva (2016).

friend's success. The motivational component, on the other hand, cannot be nil, not with anything that deserves the label 'character trait.' A purely behavioral disposition, such as the propensity to put more weight on the right side of one's body as one walks, is not a character trait, despite the fact that it may meet other criteria, such as consistency across contexts. What interests me in this paper is specifically the motivational component of traits. This is why in the initial formulation, I asked whether traits can be *based* on brute psychological facts and not whether they can *be* such facts: purely psychological attitudes may not have a behavioral component at all. For instance, one may have a more pleasurable sensation when looking at the paintings in one's dentist's office compared to the sensation one gets when looking at the paintings in one's Dean's office or vice versa, and this may not lead to any difference in behavior. So psychologically motivated attitudes cannot be expected, in general, to be co-extensive with traits. However, psychologically motivated behavioral dispositions *can* be expected to be so, or so I will argue in what follows.

Of course, in the limiting case in which a trait's behavioral component is nil and we ascribe the trait to someone on purely motivational grounds, the trait may *be* nothing other than a psychological reaction. For instance, if secretly experiencing disgust at the sight of two men kissing counts as homophobia in the absence of any behavioral manifestations,[13] then the question at issue here becomes a question about how, if at all, homophobia itself – not simply its motivational component – differs from a mere revulsion.

2. THREE POSSIBILITIES: CONTROL, ACTIVE PARTICIPATION, AND CHANGING MOTIVATION THROUGH REASONING

Let us accept the evidence from the examples I derived from Tarantino and Amelie Nothomb and cede, for the sake of argument, that the motivational component of traits *can* consist in psychological motivational facts, such as facts about immediate attractions and aversions. Perhaps there is a difference between traits and purely psychological conditions which lies not in the nature of motivation but somewhere else. Where else might such a difference lie?

One possible answer is this: the difference in question has to do not with the *nature* of motivation but with its *intensity*, and from here, with the level of agential *control*. Traits, as Sartre once suggested, incline but do not necessitate: the untrustworthy person may feel tempted to reveal a friend's secret, but she is not *bound* to do that. Perhaps things stand otherwise with psychologically motivated attitudes and dispositions. Maybe a person with a revulsion toward rats is either unable or almost unable to resist acting on it, while a person disgusted by homophobic acts is perfectly able to resist. If so, then it will follow that we have a kind of control over traits that we do not have over purely psychological conditions. And if that's right, then perhaps it is no surprise that we hold each other morally responsible for traits but not for purely psychological attitudes and propensities – more control breeds more responsibility.

But this cannot be the answer. First, it is unclear whether there are any irresistible desires.[14] More importantly, even if there are, they would be those that underlie extreme cases of pathology, not the broader class of psychological attitudes and dispositions at issue here. On the other hand, character traits may exert an extremely strong influence on behavior, and there is no reason to suppose that that influence is systematically weaker in the case of traits compared to purely psychological dispositions. For instance, a very curious person may find it extremely difficult not to pry into others' affairs, and a person with a strong sense of pride may find herself unable to grovel

---

[13] This may or may not be actually possible. It could be that we need to know a person's second order attitudes toward his or first-order reactions. For instance, if he not only conceals his disgust from others but feels very bad about it, we may intuit that he is not really homophobic, not even secretly.

[14] See Mele (2004).

even when doing so is the only way for her to save herself or her family from death. Indeed, the strength of a person's character may manifest itself *precisely* in an inability to act contrary to one's character dispositions.[15] This latter kind of motivation may be much stronger than some purely psychological propensity such as a desire to avoid the smell of gasoline. It may even be stronger than a pathological aversion to spiders (an arachnophobe may find it more difficult to grovel than to touch a spider). Finally, whatever the relative intensities, it is in principle possible for us to discover a way to control any psychologically motivated disposition. Perhaps we can develop pills that allow us to reverse the emotional valence of attractions and aversion. Intuitively, this will not turn a pure revulsion into a character trait.

This last possibility suggests another option: maybe the question is not *whether* we have control but what *kind* of control we have. Perhaps what sets traits apart from psychologically motivated dispositions is that, while we may have partial control over both the former and the latter (in the sense that we can decide whether or not to allow them to influence our behavior), the *kind* of control we exercise over traits is of a special sort: we *actively* participate in controlling our traits in a way in which we cannot actively participate in controlling our psychologically motivated dispositions. When we seek to change our characters, we do not just set a mechanism of change disconnected from reasoning in motion, as when a person takes a pill. Nor do we simply condition ourselves to respond differently in certain sets of circumstances, as we might when, for instance, we attempt to free ourselves of fear of heights through exposure to heights.[16]

But 'active participation' cannot be the answer either since, surely, we can imagine having a kind of control over psychological conditions that we exercise quite actively. This will not turn first-order aversions into traits.[17] Thus, suppose we determine that attractions and aversions can be controlled by diverting one's attention away from their objects and focusing on something else; for instance, we find out that a person averse to rats can control the aversion by focusing on the similarity between a rat's head and a squirrel's head and diverting her attention away from the rat's tail. A person averse to rats who follows this strategy may become able to *actively* control her aversion and perhaps dispose of it. This will not turn the aversion into a trait.

On the flip side, we may discover a way of moulding real traits by setting a mechanism of change in motion in a third-personal way, as when oxytocin is administered in an attempt to increase a person's level of trust.[18] This type of intervention does not put traits on a par with phobias or non-pathological psychological conditions either.

But perhaps, while 'active participation' is not precisely the distinguishing mark we are looking for, something else in this neighbourhood is: control through reasoning. As noted in the beginning, we suppose that we can, at least in principle, change traits through reasoning. We do not, in general, expect purely psychological attitudes to be responsive to reasons in the same way; for instance, we believe we can, at least in principle, make ourselves more generous people by reflecting on the reasons to be generous, for example, reflecting on the undeserved pain that a malnourished child in Somalia must endure. In contrast, we do not think we can just convince ourselves to like the smell of gasoline.

This is not the answer either, however. Consider the fact that in certain kinds of cognitive-behavioral therapy, reasoning is crucial. There, a person may be advised to reflect on the existing evidence that planes are safe and safer than automobiles so she may cure her aerophobia. This type of

---

[15] This inability is what Williams (1993) called 'moral incapacity' and Frankfurt (1988) called 'the unthinkable.' There are some differences between Williams and Frankfurt on this score. For a careful discussion of them, see Watson (2002) and Frankfurt (2002).

[16] See Marks (1979).

[17] A point recognized by Richard Moran (2002).

[18] See Kosfeld, et al. (2005).

'control through reasoning' does not turn aerophobia into a character trait either.

Note: I do not wish to suggest here that we are *guaranteed* to have cognitive control over psychological conditions, whether pathological or not. While an aerophobe can choose to carefully consider the evidence of the safety of airplanes, there is likely to be a gap between her intellectual grasp of the evidence and a change in her motivation. Reason by itself may fail to free her of fear. This gap, however, does not give us the desired clue as to the difference between traits and psychological conditions. It does not, because a parallel gap exists between the point at which we see and endorse reasons to change our character-related motivation, and the point at which our motivation actually changes. Thus, while I may make myself a continent person by persuading myself that there are reasons to behave generously even when I do not feel like it or kindly toward someone who angers me, there is a gap to be bridged between continence and the full-fledged virtues of generosity and kindness, where the latter require a change in motivation I cannot effect through reason alone. So in this respect as well, traits seem no different from pure psychological conditions. What other options are there?

3. THE PSYCHOLOGICAL FACTS VIEW AND THE NORMATIVE REASONS VIEW OF TRAITS

Another possible route to take here – perhaps the most obvious one – is to simply deny the alleged connection between traits and reasons. According to Sophia Moreau, this is what Bernard Williams proposes to do: on Williams's view, in Moreau's reconstruction of it, character traits such as loyalty are based on psychological motivational facts – the loyal person is precluded from betraying her friend by facts about her own motivation: she finds that she just *cannot* betray her friend, she has something akin to an inhibition.[19] Call this the 'psychological facts' view of traits.[20] As we shall see shortly, Williams would not accept this reconstruction of his view without qualifications, but let us set this exegetical point aside for now.

The psychological facts view has the advantage of easily accommodating examples such as psychologically motivated racism or sexism. However, it has the disadvantage of affording no resources to explain the intuitive difference we found above between character traits and dispositions that are grounded entirely in brute psychological facts. This is precisely the objection Sophia Moreau has raised apropos of the psychological facts view. Moreau demurs at what she takes to be Williams' account of loyalty as arising from some internal psychological inhibition. She writes, "'I can't do that! That would be disloyal!" does not appear to me as a brute fact, a fact about myself that I am simply landed with. Rather, I see myself as answerable for this psychological fact: I believe I could be asked to give reasons for it and that I could be in error about these reasons.'[21] According to Moreau, traits consist in dispositions to respond to normative reasons, and those reasons are grounded in normative facts existing independently of us, not in psychological facts we are 'saddled' with. This is why, in her view, we hold each other responsible for traits but not for purely psychological conditions. Call Moreau's account the 'normative reasons view' of traits.

The normative reasons view has intuitive force, which I shall endeavor to explain later. What I wish to point out here is that the view cannot ultimately succeed.[22] It flies in the face of the data. As the examples above suggest, traits can be based on psychological motivational facts, such as

---

[19] See Williams (1993).

[20] This is really a view of the motivational component of traits, but I take it that the clarification regarding the elements of traits I made at the end of Section 1 makes dwelling on such niceties obsolete.

[21] Moreau (2005, 294).

[22] I have argued for this claim in detail in Fileva (2016).

those that underlie purely psychological conditions: Tach, the SS officer, or homophobes motivated by disgust *are* sexist, racist, and homophobic, respectively, and not just people with bad taste, for instance. But then, how *do* traits differ from purely psychological conditions?

4. INPUT IN VERSUS CONCLUSION OF DELIBERATION

Williams actually directly addresses a version of this question. As I mentioned earlier, he would not fully endorse the psychological facts view of character Sophia Moreau ascribes to him, at least not without important qualifications. He would resist Moreau's objection to the effect that his proposal affords us no resources to distinguish between traits and psychological conditions. It is now time to revisit Williams's account, because Williams has a proposal about how to distinguish (an important aspect of) moral character from purely psychological conditions.

Williams addresses an aspect of the problem that is my focus here, in the context of a discussion of what he calls 'moral incapacity.' By 'moral incapacity,' Williams means the sort of incapacity that makes such statements as, 'I cannot betray my friend,' or Luther's, 'Here I stand, I can do no other,' true. Williams makes an attempt to differentiate moral incapacities thus understood from incapacities that are purely psychological, such as an inability to cross the Avon Gorge because one is afraid. This distinction is a special case of the distinction I am after, since I am interested in character and character traits more broadly. Moral incapacities can be thought of as the most central aspects of a person's character – unlike traits more generally, they tell us not simply what a person is inclined or reluctant to do, but what she, given who she is, cannot do, at least not without changing her character (in contrast, a procrastinator who meets a deadline, say, overcomes a very different sort of incapacity and cannot be said to have changed her character as a result, although the meeting of the deadline might signal that she is capable of becoming a punctual person). Yet Williams is clearly concerned with the same core problem: how to distinguish character from purely psychological aspects of the self. The suggestion he makes, if successful in solving the narrower problem that occupies him, may still give us the clue we need. What interests me now is whether the proposal succeeds. In the course of answering this question, I will also say something about the precise point of contention between Moreau and Williams.

Williams's suggestion, in brief, is the following: moral incapacities differ from purely psychological incapacities in that the latter function as *inputs* in deliberation while the former function as *conclusions*. I do not *conclude* that I am afraid to cross the Avon Gorge after considering pros and cons. Rather, I take my own fear and reluctance as input and decide what to do given that I am afraid. In contrast, that I cannot betray my friend is best thought of as a conclusion of deliberation, rather than input.

This proposal can be countered by pointing out that what Williams calls 'moral incapacities' are not always preceded by deliberation and so cannot always be conclusions of deliberations. My conclusion that I cannot betray my friend may not be anything I have ever deliberated about: I may just find myself unable to betray my friend. Williams, however, is aware of this. He insists that, nonetheless, '[t]he idea of a possible deliberation by the agent in such terms gives us the best picture of what the incapacity is.'[23] This pronouncement is a bit cryptic, but the central idea should be clear enough: moral incapacities exist only for agents who can deliberate. Purely psychological incapacities do not presuppose such an ability: a two-year old child who cannot yet make practical judgments may well have a psychological incapacity, for instance, she may be unable to get out of bed in the dark because of fear. But a two-year old child cannot be said to have a moral incapacity of any sort.

Can Williams's proposal succeed? The answer is 'no,' but in order to see why, we must go back to Moreau's interpretation of Williams. Moreau, as we saw, accuses Williams of espousing a

---

[23] Williams (1993, 65).

psychological facts view of character on which character traits are not properly connected to reasons and reasoning. Yet, I just said that for Williams, character traits such as those that constitute moral incapacities are best thought of as conclusions of deliberation. But conclusions of deliberation are intimately connected to reasons. Is Moreau wrong in her interpretation of Williams, then?

Given Moreau's own view of reasons, the answer to this question is 'no' as well. While Williams connects the sorts of moral incapacities constitutive of someone's character to reasons and reasoning, he has in mind reasons *as he understands this term*. His view is one on which all reasons are 'internal' in the sense that something is a reason for an agent only if she can see it as a reason *given her motivation*. The last point is crucial: character is not grounded in external reasons, understood as reasons that apply to an agent independently of her motivation. According to Moreau, grounding character in internal reasons effectively amounts to equating character traits with psychological conditions, since in the absence of external reasons, character becomes purely a matter of psychological motivational facts about the agent.

Moreau, as I will now suggest, is correct in the following sense: if there are no normative reasons applicable to us independently of our motivation, character traits and essential aspects of character such as moral incapacities cannot be distinguished from purely psychological conditions. The twist is that Moreau's own account of character traits fails: character traits need not be *motivated* by reasons in any straightforward sense. The motivational component of traits can consist in psychological motivational facts about their possessor. How, then, do traits differ from psychological conditions?

5. THE REAL DIFFERENCE: APPLICABILITY OF NORMATIVE REASONS

I will begin by asking the reader to consider a thought experiment. Suppose that rats, as a result of various scientific experiments, evolve as a species and develop moral and emotional capacities. They acquire the emotional capacity to suffer humiliation much the way humans can. If that happens, we would have to change our attitude toward rats: we would have to respect them and avoid running away from them. Fear of or revulsion toward rats will cease to be purely psychological conditions and will morph into a morally-laden trait. But nothing about the psychological origin of that fear or that revulsion need have changed. The aversion may continue to arise as a result of an immediate gut reaction. This suggests that the distinction between traits and psychologically motivated attitudes hinges on the applicability of normative reasons. Racism is a trait because there is a normative reason not to be racist. This reason has to do with the value of humanity. It applies to a person even if he or she does not recognize its existence. The point is not that the best picture of, say, white racism, of what it means to say, 'I cannot treat non-whites as being on a par with whites,' is the picture that presents this assertion as a conclusion of deliberation, as per Williams's interpretation of moral incapacity, since this is clearly *not* the best picture. Nor is it that racism must be motivated by reasons – while it certainly *can* be, as when it is based on a misguided ideology, it does not *have* to be. The point is, rather, that an aversion-based propensity with a psychological origin may or may not be a character trait, depending on its object and the moral reasons or lack thereof for treating that object in a certain way. This is because when a person's dispositions are subject to moral standards, they can be seen as expressive of his or her moral values. Homophobia and racism are traits while rat-repulsion is not because the former two but not the latter can be viewed as part of our normative value system.[24]

---

[24] Of course, inflicting pain on rats – as a cruel juvenile might – could be a sign of a very serious character flaw. My point here is simply that merely avoiding rats is not such a flaw, though a failure to respect rats and a tendency to systematically avoid them would be a flaw (akin to racism) were rats to acquire moral capacities.

But what makes it possible for normative reasons to apply to us? Surely, if the motivational component of traits consisted in immutable and recalcitrant psychological facts, it would be unfair to normatively evaluate traits. This point is well taken. Normative reasons are only applicable to us because we have capacities for reflection and control.[25] A person such as the SS officer or the writer Tach from the earlier example, if they detect in themselves hatred of an unknown origin against a group of people, should, at the very least, abstain from acting on it and, preferably, attempt to change it. But it would not be true that they should change it if they cannot change it. Thus, there is truth to the suggestions discussed earlier, that character traits are connected to our ability to actively participate in shaping them. What is *not* true is that these capacities can help us differentiate character traits from psychological conditions: as we saw, we can actively participate in the moulding of psychological propensities, and we can even exercise control through reasoning, much the way we can participate in the moulding of our traits; it is just that when it comes to psychological propensities, for the most part, we have no reason to.

Note also that when normative reasons do not apply, extreme fragmentation becomes possible. This is particularly clear in the case of pathological psychological conditions such as phobias. Thus, the bravest person may be arachnophobic. A person could be *truly* brave and *truly* an arachnophobe. In contrast, a person cannot *truly* value humanity and *be racist* at the same time. Someone who values all of humanity except one race does not, in fact, truly value humanity.

Of course, traits can be fragmented to one or another extent. But extreme fragmentation makes us uncertain who someone is, what her values and traits are.[26] Typically, there is a certain unity to the aspects of someone's character, and there is a way in which these aspects 'fit together.' To 'know a person well' means, in part at least, precisely to have a sense of how a person's values are connected.[27]

I must point out, however, that here, as elsewhere, there is a useful distinction to be drawn, but there is no sharp boundary. Traits are not a unified category and some are closer to psychological conditions than others are. In particular, traits reflective of a first-order propensity that does not harmonize with a person's considered preferences and value rankings resemble psychological conditions, for instance, curiosity in a person who would much rather be focused on her own life, and who may even feel humiliated by her own curiosity, thinking that there is something painful in the fact that she takes so much interest in the lives of others who take no interest in hers.

---

[25] I argue elsewhere that the reflective capacities of psychopaths are impaired, and that psychopathy is, therefore, truly a medical condition rather than a cluster of moral vices. See Fileva (n.d.).

[26] And it would normally signify a pathology, as in the case of the condition known as Dissociative Identity Disorder (DID).

[27] It can be objected here that people's characters do not have the kind of unity I just claimed they do have, because our behaviors are not cross-situationally consistent. This is an old objection, which goes back to Hartshorne and May's (1928) study on honesty in children, which marked the beginning of what is now known as the person-situation debate in psychology (from which the philosophical debate over the empirical tenability of virtue ethics springs). Allport (1937) offered a masterful response to the early situationists. Allport's defence of personality theory convinced his contemporaries, and the debate ceased for about three decades. It was reignited in the late 1960s, with the publication of Mischel's (1968). Psychologists are currently in a third, reconciliatory phase of the person-situation debate. For Mischel's current position, see Mischel (2009) and Mischel, Shoda, & Denton (2002). For a promising reconciliatory project, see Fleeson (2004). Fleeson argues that when it comes to performance on a single occasion, situational variables have more predictive power, whereas when it comes to behavior over time, personality variables do. In philosophy, sceptical views of character have been defended by John Doris (2002) and Gilbert Harman (1999). Harman and Doris drew on social psychology, primarily on the work of Ross and Nisbett (1980; 1991). For some good responses by philosophers to character scepticism, see Kupperman (2001), Kamtekar (2004), and Sreenivasan (2002).

6. PSYCHOLOGIZING VICE BUT NOT VIRTUE?

There is a final possibility I wish to consider. We tend to see the motivational component of character traits differently depending on whether we are thinking of virtues or of vices. Moreau's normative reasons view of traits may seem to capture the essence of our folk ideas about virtue even if it does not capture our intuitions about vice. It may seem that a loyal person cannot see her own loyalty as a pure psychological inhibition. If someone wanted to betray his friend but could not do it due to such an inhibition, we may be inclined to say that the person's incapacity is no longer moral and has become purely psychological.[28] If this is right, then the normative reasons view may give us the correct account of virtue without giving us the correct account of vice.

However, it is doubtful that this is right. It is true that examples of psychologically motivated vices are easier to come by than are examples of psychologically motivated virtues – it is no coincidence that all examples I have discussed here have been examples of vices. Yet, the motivational component of virtues may have a more tenuous connection to reasons than the proposal suggests. Thus, while Moreau accuses Williams of failing to properly incorporate the role of reasons and reasoning in his account of character, another one of Williams's critics, Craig Taylor, attacks Williams from the opposite side, arguing that there are cases of moral incapacity – and so, paradigmatic cases of character – in which deliberation plays no role, not even in Williams's (admittedly attenuated) sense. Taylor gives as an example Huckleberry Finn: Huck Finn's incapacity to give up Jim is, intuitively, moral and a matter of character, but it is not preceded by deliberation or otherwise grounded in Huck's deliberative process; quite the contrary, when he deliberates, he concludes that he ought to give up Jim. Huck perceives his own inability to act on his deliberative conclusion as a kind of character weakness.[29]

Of course, one can argue here that Huck Finn acts for reasons, just not for reasons he is aware of – he acts, one might say, on unconscious reasons.[30] In some sense this is right, but not in a sense that would make a normative reasons view of virtue correct. Huck Finn's inability to give up Jim, I maintain, is grounded in unendorsed attachment to Jim. There is virtue in this inability, but the virtue is connected to normative reasons only in the sense that normative reasons *apply* to Huck: there are strong moral reasons to act just as he does. And Huck is, in principle, capable of grasping those reasons. There is, however, no evidence that he, in fact, acts on those reasons.

7. CONCLUSION

In the beginning of this essay, I asked two questions: first, can the motivational component of traits consist in brute psychological facts; and second, on the assumption that the answer to the first question is 'yes,' how can traits be distinguished from purely psychological attitudes and dispositions, such as phobias and aversions. I argued, on the basis of examples, that the first question must be answered in the affirmative. I then discussed a number of ways to distinguish between character traits and psychologically motivated attitudes and dispositions, and I claimed

---

[28] Harry Frankfurt (1988) makes a suggestion along these lines.

[29] Taylor (2001) and his earlier (1995). For a response sympathetic to Williams and critical of Taylor, see Clark (1999).

[30] At this point, my discussion intersects with the debate over whether moral judgments are typically based on reasons. Jonathan Haidt (2001) makes the related point that by and large, moral judgments are based on gut reactions. Daniel Jacobson (2012) responds by arguing that Haidt confuses people's inability to give reasons for their judgments with the absence of reasons. See also Railton (2014). There are two points to note here: first, Jacobson's interpretation appears plausible only if, *in fact*, there are good reasons to make a certain judgment. Second, even in those cases in which the claim seems plausible, we need an argument in order to accept it as true: an argument to the effect that if there are reasons, then motivation seemingly stemming from psychological facts really stems from those reasons, while in the absence of reasons, motivation with the same phenomenology is what it appears to be – based on brute psychological facts.

that, ultimately, the difference has to do not with the origin of motivation but with the applicability of normative reasons. I wish to note, however, that the normative reasons that apply to us, together with our reflective capacities, influence motivation. The SS officer's racism is of psychological origin, but it is not *entirely* based on psychological facts: it is based also, partly, in the officer's tacit endorsement, whether active or passive, in his 'going along' with the psychological inclination, in his unwillingness to respond to the relevant normative reasons. Thus, while the difference between character traits and psychological conditions does not have primarily to do with motivation, this thesis must be qualified: the applicable normative reasons change the character of an agent's motivation, even though they may not change its psychological origin. Reasons can turn what would otherwise be a brute psychological disposition into a character trait. Note that these considerations do not imply an 'endorsement' criterion of character. We can possess a character trait without endorsing it even tacitly and may continue to have it despite actively fighting against it. For instance, a person may be envious without wanting to be, although if he controls his behavior, he may not have the full-fledged version of the vice of envy. What is not possible is to have the vice of envy in the absence of reasons not to be envious.

ACKNOWLEDGMENTS

References

Allport, G. (1937). *Personality: A Psychological Interpretation*. New York: Holt, Rinehart, & Winston.

Aristotle. (2003/349 B.C.) *Nicomachean Ethics*, Tredennick, H. (ed.). New York: Penguin Classics.

Belkin, L. (1988). Texas judge eases sentence for killer of two homosexuals. *New York Times*, December 17.

Clark J. & Fessler, D. (2015). The role of norms in disgust and of disgust in norms research: why liberals shouldn't be morally disgusted by moral disgust. *Topoi*, 34, 483–498.

Clark, M. (1999). Moral incapacity and deliberation. *Ratio*, 12, 1–13.

Commonwealth v Carr. (1990). 580 A.2d 1362 (Pa Supr. Ct.).

Deigh, J. (2006). The politics of disgust and shame. *Journal of Applied Philosophy*, 10, 383–418.

Dodds, E. (1945). Plato and the irrational. *The Journal of Hellenic Studies*, 65, 16–25.

Doris, J. (2002). *Lack of Character: Personality and Moral Behavior*. New York: Cambridge University Press.

Fileva, I. (2016). 'Two senses of "why": traits and reasons in the explanation of action', Fileva, I. (ed.), *Questions of Character*. New York: Oxford University Press.

———. (n.d.). 'Reflection without empathy: the case of psychopathy', in progress.

Fleeson, W. (2004). Moving personality beyond the person-situation debate. *Current Directions in Psychological Science*, 13, 83–87.

Frankfurt, H. (1988). 'Rationality and the Unthinkable', *The Importance of What We Care About*. Cambridge: Cambridge University Press.

———. (2002). 'Reply to Watson', Buss, S. & Overton, L. (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt*. Boston, MA: MIT Press.

Gooch, P. (1971). 'Vice is ignorance': the interpretation of *Sophist* 226A–231B. *Phoenix*, 25, 124–133.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834.

Harman, G. (1999). Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, 99, 315–331.

Hartshorne, H. & May, M. (1928). *Studies in the Nature of Character*, vol. 1, *Studies in Deceit*. New York: Macmillan.

Hume, D. (1975/1748). *An Enquiry Concerning the Principles of Morals* in Selby-Bigge, L.A. (ed.), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Oxford: Clarendon Press.

Jacobson, D. (2012). 'Moral Dumbfounding and Moral Stupefaction', Timmons, M. (ed.), *Oxford Studies in Normative Ethics*. Oxford: Oxford University Press.

Kahan, D. (1988). The anatomy of disgust in criminal law. *Michigan Law Review*, 96, 1621–1657.

Kamtekar, R. (2004). Situationism and virtue ethics on the content of our character. *Ethics*, 114, 458–491.

Kant, I. (1991/1797). *The Metaphysics of Morals*, Gregor, M. (tr.). Cambridge: Cambridge University Press.

Kosfeld, M., Heinrichs, M., Zak, P., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435, 673–676.

Kupperman, J. (2001). The indispensability of character. *Philosophy*, 76, 239–250.

Marks, I. (1979). Exposure therapy for phobias and obsessive-compulsive disorders. *Hospital Practice*, 14, 101–108.

McDowell, J. (1979). Virtue and reason. *The Monist*, 62, 331–350.

Mele, A. (2004). 'Volitional disorder and addiction', Radden, J. (ed.), *The Philosophy of Psychiatry: A Companion*. Oxford: Oxford University Press.

Miller, W. I. (1997). *The Anatomy of Disgust*. Cambridge, MA: Harvard University Press.

Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.

———. (2009). From *Personality and Assessment* (1968) to personality science, 2009. *Journal of Research in Personality*, 43, 282–290.

Mischel, W., Shoda, Y., & Denton, R. (2002). Situation-behavior profiles as a locus of consistency in personality. *Current Directions in Psychological Science*, 11, 50–54.

Moran, R. (2002). 'Frankfurt on identification: ambiguities of activity in mental life', Buss, S. & Overton, L. (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt*. Boston, MA: MIT Press.

Moreau, S. (2005). Reasons and character. *Ethics*, 115, 272–305.

Murdoch, I. (2001). *The Sovereignty of the Good*. Abington, UK: Routledge.

Nothomb, A. (2010). *Hygiene and the Assassin*, Alison Anderson (tr.). New York: Europa Editions.

Nussbaum, M. (2010). *From Disgust to Humanity*. New York: Oxford University Press.

O'Brien, M. (1967). *The Socratic Paradoxes and the Greek Mind*. Chapel Hill, NC: University of North Carolina Press.

Plakias, A. (2013). The good and the gross. *Ethical Theory and Moral Practice*, 16, 261–278.

Railton, P. (2014). The affective dog and its rational tale. *Ethics*, 124, 813–859.

Ross, L. & Nisbett, R. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

———. (1991). *The Person and the Situation: Perspectives on Social Psychology*. New York: McGraw-Hill.

Sreenivasan, G. (2002). Errors about errors: virtue theory and trait attribution. *Mind*, 111, 47–68.

Taylor, C. (1995). Moral incapacity. *Philosophy*, 70, 272–285.

———. (2001). Moral incapacity and Huckleberry Finn. *Ratio*, 14, 56–67.

Watson, G. (2002). 'Volitional Necessities', Buss, S. & Overton, L. (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt*. Boston, MA: MIT Press.

Williams, B. (1993). 'Moral incapacity,' *Proceedings of the Aristotelian Society*, 93, 59–70.