PUBLIC REASON

Journal of Political and Moral Philosophy

Volume 14, Number 2, 2022 & Volume 15, Number 1, 2023

Editura Universității din București

PUBLIC REASON

Journal of Political and Moral Philosophy

Public Reason is a peer-reviewed journal of political and moral philosophy. Public Reason publishes articles, book reviews, as well as discussion notes from all the fields of political philosophy and ethics, including political theory, applied ethics, and legal philosophy. Texts from other areas of philosophy may be published, if they are relevant for issues in moral and political philosophy. The journal encourages the debate around rationality in politics and ethics in the larger context of the discussion concerning rationality as a philosophical problem. Public Reason is committed to a pluralistic approach, promoting interdisciplinary and original perspectives as long as the ideals of critical thinking and clarity are respected. The journal is intended for the international philosophical community, as well as for a broader public interested in political and moral philosophy. Public Reason publishes two issues per year, in June and December. Public Reason is an open-access e-journal, but it is also available in print. Public Reason is an open-access journal, which does not charge authors a processing fee – neither directly nor indirectly (where applicable) through the author's institution of affiliation. Hence, publication-related decisions are not made on financial considerations

Public Reason is available online at **http://publicreason.ro**
ISSN 2065-7285
EISSN 2065-8958

PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 14, No. 2, 2022 & Vol. 15, No. 1, 2023

ARTICLES

Public Reason 14(2)-15(1): 3-18

On Vindication in Ethical Life[1]

David Owen
University of Southampton

Abstract. This paper introduces reflection on the concept of vindication in ethical life. It distinguishes vindication from justification and asks how the reflective stances towards our ethical agency that they institute differ. It argues that these reflective stances, and the forms of accounting that each institutes are importantly distinct but also necessary elements of any ethical outlook, and ones that can be related to one another in more than one way. It concludes by relating this argument to themes from the 'morality critics', especially those of meaning in life and moral luck raised by Susan Wolf and Bernard Williams.

Keywords: vindication, justification, moral luck, meaning in life, challenge, pride, joy.

One question that can be asked of us, as ethical agents, is whether our actions are justified; a second question is whether our actions are vindicated by what, it turns out, we have done. Each of these questions constructs a reflective stance towards our ethical agency that involves giving an account of ourselves. But what are the features of these reflective stances? How, if at all, are the accounts each calls for different? In what kind of relationship do they stand? My aim in what follows is to address these questions. I will argue that these reflective stances, and the forms of accounting that each institutes, are importantly distinct but also necessary elements of any ethical outlook, and ones that can be related to one another in more than one way. I will conclude by relating this argument to themes from the 'morality critics', especially those of meaning in life and moral luck raised by Susan Wolf and Bernard Williams.

1] The ideas in this paper have been meandering around my mind for more years than would forgive the still provisional form of their expression in this paper. Important spurs were provided by reading Bernard Reginster's *The Affirmation of Life* (2006) and R. Jay Wallace's *The View From Here* (2013) and a first occasion was provided by an invitation from Jamie Draper to give a keynote address to the REAPP conference at the University of Reading in 2018. Since then various inchoate versions and spin-offs of this material have been presented at the Universities of Cambridge, Frankfurt, Leiden, Leuven, and Southampton as well as at IAS Princeton during a period as visiting professor there in 2021 and to the online Bernard Williams group. I am grateful to all these audiences for their efforts to help me clarify and develop my thoughts, however limited their success may have been. More particular thanks for encouragement, comments and objections are owed to Chris Armstrong, Duncan Bell, Sophie Grace Chappell, Didier Fassin, Rainer Forst, Thomas Fossen, Mike Gadomski, Anca Gheaus, Jonathan Havercroft, Tim Heysse, Cindy Holder, Paul Katsafanas, Avery Kolers, Matthew Kramer, Nikhil Krishan, Dorota Mokrosinska, Geraldine Ng, Matthieu Queloz, Ben Saunders, Joerg Schaub, Francesco Testini, Alan Thomas and Jake Wojtowicz as well as, most especially, my friend and colleague Tracy Strong who left us too early. The current version of this argument was presented at the Rousseau Lecture at ECPR House in December 2022. I am very grateful to Sorin Baiasu for the invitation and the interlocutors and commentators who participated in the symposium that followed – Matt Bennett, Allan Gonzalez, Sacha Mudd, Paula Satne, Joe Saunders, Jakub Szczepanski, Francesco Testini and Previn Karian.

This argument begins by sketching the contours of the concept of vindication through consideration of a series of examples. It then turns to lay out what it proposes are the forms of reflective stance towards our ethical agency that the concepts of justification and of vindication establish – and argues that these involve two distinct ideals of ethical life. In the third part of the paper, I aim to illustrate and support the claim that the forms of accounting (and ideals) that each of these reflective stances institute are necessary elements of any ethical outlook and that they can stand in diverse ways with respect to each other by distinguishing three different kinds of relationship that they can exhibit.

I. CONTOURS OF THE CONCEPT OF VINDICATION

When we think about vindication in the ordinary everyday sense of the word, we may have in mind cases like that presented by Terence Rattigan's play *The Winslow Boy* (set in the Edwardian era and loosely based on real events) in which a 14 year old boy (Ronnie) is accused of the theft of a postal order and expelled from the naval college he attends. The consequences of this event for his family's honour and the boy's future are significant. The family would be shunned by decent society and the boy marked by a permanent stain on his character. Convinced of his innocence, on the basis of his word alone, the family then engage in a long, uncertain and expensive quest for justice which demands sacrifices from them all in terms of the father's health, the brother's Oxford tuition and his future career in the civil service, and the daughter's marital options. They are eventually successful, and to their joy Ronnie's name is cleared. They affirm the value of their having acted as they did in the light of the value of the outcome that they have successfully realised; indeed, they can take pride in the outcome as something that they have brought about: "We did it!"

It is easy in the light of this kind of example to think of vindication as concerning successful struggles to overcome injustice or wrong; but while it may involve such features, it need not do so as a second example illustrates.

In 1984 Nick Faldo was a leading European golfer, having won five tournaments the previous year, but he had not seriously challenged in the majors despite getting into positions to do so. Convinced that his overall ability was sufficient to contest majors, Faldo embarked on the radical and highly risky strategy of completely rebuilding the long swing that he saw as his key weak point. As Faldo's golf began to deteriorate at alarming speed, a number of his sponsors jumped ship. Faldo's livelihood was becoming more and more precarious. He did not win another tournament for three years, but in 1987, to his joy, he won the first of what would turn out to be six major championships. Faldo's decision to rebuild his swing was vindicated and he was entitled to take pride in the fact that he did it (something he expressed by performing a rather tuneless quotation from the song "My Way" at his first major triumph). But notice that his course of action

would have been vindicated even if he had won fewer majors or even none at all (given the role that extrinsic luck plays in such events), as long as it was the case that he was able to seriously challenge in them and did not lose them or gift the win to another competitor (i.e., that his technique and nerve did not collapse under pressure as, for example, Greg Norman's did in the 1996 US Master's when he blew a six shot lead over Faldo in the final round).

The key elements across these two examples are that the course of action involved struggle (that the outcome was uncertain and that it required sacrifice for its achievement) and that the process of struggle realizes outcomes of value to the flourishing of the agent's life such that they have, all things considered, reason to affirm - or at least not regret - that they have acted as they have.[2] Whereas the outcome of value in the first example is the realization of an external goal (Ronnie's name is cleared), in the second example it is the realization of a transformation of the agent's powers (Faldo's ability to maintain a consistent swing under intense competitive pressure). If we imagine for a moment either that these outcomes could have been achieved by the relevant agents simply by wishing for them or that their actions had no relationship to the outcome realised, the concept of ethical vindication would have no application because there would be no achievement in which to feel joy or take pride. Perhaps it needs to be stressed here that these actors do not engage in their actions in order to experience joy or feel pride, rather they engage in their actions because what they are trying to realize is valuable to them, because they care about these matters.

It is important though not to be misled by two shared features of these initial examples, namely, that in both cases the value of the outcome is specified by a goal that is given in advance of the agent's conduct and the outcome is successfully realised. Neither of these are necessary for the exercise of one's agency to be vindicated.

Firstly, it can be the case that the goal is transformed in the process of struggle and the new goal is achieved. One example of this phenomenon is Tchaikovsky's Piano Concerto No. 3 in E-flat major which was initially conceived by him as a symphony in the same key, however his pursuit of this goal led him to realise that what he wanted to express did not work in this form and he abandoned it, reworking the first movement as a one-movement *Allegro brillante* for piano and orchestra. Another example would be that of a couple who commit to marriage and come to see through their joint efforts to

2] Does vindication require struggle? Matt Kramer has proposed the following counter-example to me: 'Suppose that I've been spending five pounds each week on lottery tickets and that I eventually desist from that pattern of behavior. Suppose that I then decide to resume my weekly purchases of lottery tickets, and a few weeks later I win the lottery. I can warrantedly feel vindicated in my decision to resume that pattern of behavior even though I have not engaged in any struggle. My decision was of course reached on the basis of a high degree of uncertainty, but the resultant course of action involved only a trifling sacrifice.' This seems plausible but I think it lacks the right kind of relationship of agency and outcome to be constitutively part of the flourishing of one's life.

realise a good marriage that they are best suited to live apart as good friends rather than together as spouses and part amicably.

Secondly, it can also be the case that although the agent's pursuit of their intended goal falls short, the process of pursuing it produces other outcomes of value that are sufficient to provide reasons not to regret the course of action undertaken. An example, which I owe to Adams (2010), is Claus von Stauffenberg's attempt to rescue Germany from Nazi rule by assassinating Hitler and leading a coup d'état on July 20th 1944. This attempt failed and led not only to Stauffenberg's death but also to that of hundreds of others who had collaborated or were suspected of collaborating with him. On Adams' reconstruction, Stauffenberg (and his co-conspirators in this project) 'seem to have been motivated in large part by the meaning that they hoped their deeds would have for others, believing that even if it couldn't succeed in its own terms, the plot against Hitler should be attempted for the honor of Germany, to show the world that some Germans stood up against Hitler's crimes.' (Adams 2010, 80) Although they failed to assassinate Hitler and bring about a coup d'état, their attempt to do so did succeed in showing the world that some Germans were willing to risk their lives to stand up against the regime. If Adams' supposition concerning their motivations is correct, then although they have reason to regret that they were not successful in realizing their intended goal, they need not have reason to regret undertaking the course of action on which they embarked (despite the very high price it demanded of them). They need not because their actions are expressive of the values that give meaning to their lives and, in performing actions expressive of these values, they do realize something that matters on the terms that motivated the external goal they have failed to realise.[3] If, all things considered, they have reason to affirm their actions, they can take pride in their conduct.

These last examples, apart from illustrating the point that vindication need not be restricted to cases of successfully realising a goal that can be stated in advance of the action, should also draw our attention to an issue that we have not yet addressed. Thus far we have seen that the concept of ethical vindication has application when the agent in question has reason to affirm, on the basis of their values, the outcome to which their actions give rise. The entitlement of the agent to feeling vindicated is conditional on its being true that their actions gave rise to this outcome, but it also matters that the values that are realised in such outcomes are values that there is good reason for human beings under the situated circumstances of their lives to value ethically (even if they are not the values to which we ourselves are, or even can be,[4] committed). An action may fail to be vindicated either because the outcomes to which it gives rise are not ones which the agent has reason to affirm on the basis of the values to which they are committed *or because there*

3] I use the weaker formulation 'need not' here because a full account of whether they have reason to consider their actions vindicated would involve a wider consideration of, for example, they were negligent in, for example, their planning or coordination, and hence of the reasons for their failure to realise their external goal.

4] We can recognize the value of forms of life that are not available to us under the conditions of our lives.

is no reason to value the "values" on the basis of which the agent affirms the outcome. Thus, if we consider the Stauffenberg case, it matters here that we have reason to recognize the value of patriotism of the kind exhibited in this example[5] as a value to which ethical agents can reasonably be committed, as something that can be recognized as serving to give meaning and value to the lives of relevantly situated human beings. This does not require that we share their commitment to this value any more than acknowledging the value of works of art such as Tchaikovsky's Piano Concerto No. 3 demands that we share the aficionados' love of music. We may ourselves be largely indifferent to the value of both patriotism and music. It does, however, require that we have reason to recognise these values as values to which agents may intelligibly choose to commit themselves as ethical agents. Iago may see his actions as vindicated by the tragic outcome for Othello and Desdemona that he brings about since this outcome is expressive of his (evil) values, but we have no reason to endorse, and every reason to reject, his claim.[6]

To support this claim, we can remind ourselves that it is a feature of an action being apt for ethical vindication that it involves struggle (in the sense of overcoming resistances) and that one's agency stands in the right kind of relationship to the outcome. It is these features that allow agents to take pride or experience joy in what they have achieved and to express this in spontaneous avowals such as "I did it!". However, to take pride or joy in an achievement is to be committed to the view that the value of the outcome one has brought about is one that can and should be acknowledged by others as valuable, where 'others' here refers to humanity as transhistorical community. If we take pride in outcomes that our actions have brought about that are seen by our contemporaries as trivial or shameful, sustaining our self-assessment requires that we are committed to the claim that our contemporaries are, in some way or other, mistaken and that we can provide reasons that are intelligible to them as reasons for why they are mistaken. (Put another way, we call for them to undertake a re-evaluation of their values).

In summary, then, the concept of vindication refers to contexts of agency in which an agent engages in struggle to realise a goal (which may change though the process of struggle) and results in outcomes that give the agent reason to affirm (or at least not regret) their actions on the basis of their value-commitments and where there is general reason to recognise these outcomes as ethically valuable.

5] Avery Kolers and Cindy Holder have helpfully posed the question of whether the case of Robert E. Lee relevantly parallels the Stauffenberg example in some respects despite being directed at patriotically sustaining a regime built on slavery. Here the issue hangs on whether an agent in Lee's circumstances could reasonably be committed to the view that sustaining the institution of slavery was ethically admirable rather than utterly shameful. I don't think this is plausible.

6] As Matt Kramer rightly commented: 'A serial rapist might look back upon his spree of crimes with a feeling of vindication for his decision to embark upon them. He values the feelings of elation and mastery and sexual pleasure which he experienced through that spree. His pursuit of those feelings is coherent and intelligible, but it's obviously not ethically worthy. He ought not to feel any sense of vindication. Quite the contrary.'

REFLECTIVE STANCES: JUSTIFICATORY AND VINDICATORY

With this initial delineation of the concept of vindication in place, let us turn to the second task, that of demonstrating that the concepts of justification and of vindication make available distinct reflective stances towards our ethical agency. I will start with a provisional sketch of what is in play when we are asked to justify our actions or to vindicate our conduct, where the purpose of this sketch is to show that the questions of justification and of vindication involve different types of reflective stance towards our ethical agency that call for different kinds of accountings of ourselves, and this difference points to the distinct roles that each may play in our ethical lives.

One way to think about the concept of moral justification is to see it as responding to what we might call 'the circumstances of social morality' by which phrase I mean to highlight two features that, taken jointly, make morality both possible and necessary. The first of these is that we understand ourselves as agents who are capable of practical deliberation concerning what we should do and, at least much of the time, of acting on the outcomes of our deliberations. The second is that we are social beings whose actions can affect the conditions of each other's agency in ways that are consequential for our ability to pursue our own projects and for the maintenance of the conditions of social life. Taken together, these general features of human life both make possible and require morality as articulating 'a system of socially sanctioned demands' (what we owe to each other) that are 'also at least in some degree generally acknowledged as claims by those subject to them' and thus serve as inputs into our practical deliberations about what we should do (Strawson, 1974: 45).

The question of whether our actions are morally justified can then be seen as instituting a reflective stance and form of accountability directed at our ethical lives as social beings who owe consideration to each other. It requires that an agent demonstrate that the practical deliberations that underwrite their actions acknowledge the claims of what we owe to each other expressed in 'a system of socially sanctioned demands' and, if not, to provide reasons that (fully or partially) excuse this failure. It is important to notice that it is integral to this kind of accountability that one should be in a position to apply the relevant deliberative considerations at the time of acting: it is, as Williams remarks, 'thought to be essential … to the notion of justification itself … that one should be in a position to apply the justifying considerations at the time of choice' (Williams 1981, 24).

To offer this kind of accounting of one's actions is, thus, to demonstrate whether the exercise of our agency is (or is not) 'above board'; where justification puts the agent 'in the clear' in the sense that no complaint or claim can be upheld against them in respect of their performance of the action concerned, and hence no sanction (formal or informal) can be legitimately applied to them in virtue of this conduct. This is, then, a form of accountability directed at our moral competence that underwrites a form of

self-valuing in which the agent is judged relative to an ideal of upright conduct, one who discharges the debts they owe to others.

What of the concept of vindication and the reflective stance towards our ethical agency that it makes available to us? One way of thinking about the concept of vindication is to see it as being responsive to two basic features of our ethical lives. The first is that our agency is characterised by the pursuit of valued goals through the exercise of limited, fallible abilities under circumstances that we do not control and cannot fully foresee. The second is that the outcomes of our actions (intentional or unintentional) matter for how our life goes, its flourishing, and whether we can affirm it.

The reflective stance that this concept would seem to make available is one in which the accounting that is asked of us concerns whether the value of our doing what we have done enables us to affirm our lives as ethical agents. But to flesh out this stance we need to recall the point that our experience of the value of our agency given natural expression in the cry "I did it!" is tied to overcoming resistances to the achievement of our valued ends. To will the goal is, in this sense, to will the resistances to the achievement of the goal. Suppose, for example, Faldo had been able to transform his swing simply by wishing it, that it had not needed requires the work, sacrifice and risk that was necessary for the realization of this end. Under such conditions, Faldo would not have enjoyed the sense of achievement that underwrites his ability to affirm the value of this course of conduct to the flourishing of his life.

It can sound paradoxical to say that willing the goal means willing resistances to the goal,[7] but this appearance of paradox is dissolved when, recalling that we are limited, fallible agents acting under circumstances we don't control and cannot fully foresee, we conceive of valuable goals as challenges. We can help draw out the significance of this by considering four features of challenges:

1) they involve overcoming resistances (no resistance, no challenge);

2) they must be possible for the particular agent, although whether they are possible may not be determinable prior to the attempt (if there is no practical possibility of you achieving X, then X is not a challenge);

3) their value is at least partially related to their difficulty (given two challenges distinguished only by their degree of difficulty, the more challenging option is the more valuable);

4) once a challenge is met (if it is the kind of challenge that can be fully and finally met), then, barring significant changes in the circumstances of one's agency, it is no longer valuable as a challenge (which is not to say it has no value).

The kind of accountability that the vindicatory stance involves thus concerns whether the agent is exercising, developing, and realising their abilities to achieve

7] I am influenced here by the account of will to power in Reginster 2006 as I reconstruct it in my review of his book (Owen, 2009).

valuable ends, whether they are pursuing what matters to them and in doing so realizing valuable goals that allow them to experience themselves as the authors of lives that they can affirm or, at least, all things considered, have reason not to regret.

However, the life-affirmation I may enjoy in the successful achievement of a goal ('Hurrah, I did it at last!') has as its flipside the possibility of regret in a variety of agent-related forms that range from regret at my failure to take up opportunities to pursue particular goals to regret at the disvaluable outcomes to which my conduct inadvertently gives rise (and the attendant reflections in the form of "if only I had/had not" can haunt a life to destruction). Note further that because vindicatory judgments are always retrospective (they pertain to what, it turns out, we have done), they are situated in the wider biographical narrative of one's life, and their character and significance can change as that life-history develops. In this respect, specific vindicatory judgments are always provisional, what strikes me now as a (perhaps shameful) failure to overcome a challenge may acquire value in virtue of being an experience that I come to see as a necessary element of the successful achievement of another goal, in which case the earlier failure may be (at least partially) redeemed. Conversely, as the example of Oedipus illustrates, successful achievements of particular goals may lead to an overall outcome of a life that is tragic (i.e, cannot be vindicated). Hence the remark attributed to Solon by Herodotus: 'call no man happy until he is dead' or more fully '"Yet keep yourself from calling him happy before he dies; he is lucky."

So whereas the kind of accountability demanded from the justificatory stance is raised from the standpoint of the other to whom one owes moral consideration, the kind of accountability called for from the vindicatory stance is raised from the standpoint of one's attainable self, the future self who is liable to be haunted by the thought 'if only'. We may note that one reason that the figure of the older and wiser friend (e.g., Socrates) is a classic topos in philosophy is that they may exemplify this future standpoint in the present. The ideal of righteousness that emerges from the justificatory stance may then be contrasted to the ideal of life-affirmation that emerges from the vindicatory stance.[8]

All this is, in one respect, just a way of spelling out the claim that any ethical outlook involves both a form of accountability that concerns an agent's responsibility to others and one that addresses their responsibility to themselves (most basically to the conditions of their being able to affirm, or at least not regret, their lives). The contrast in its negative mode is, if you like, between a wrongful life and a wasted life (in the sense of 'wasted' that connects both to lack of development of one's potential and to meaninglessness). Roughly speaking, the justificatory stance recognizes that we are agents who have to work out ways of living together on terms each has reason to accept and the vindicatory stance acknowledges that each human being has a fundamental

8] This is not to say that the exercise of agency in the ways sketched is the only source of life-affirmation; on the contrary, joy in life can arise in more passive or receptive ways from encounters with art or nature. Thanks to Joerg Schaub and Paul Katsafanas for pushing me on this point.

interest in being able to see their life as meaningful and valuable in a way that connects to how they conduct it under the conditions of fortune that they encounter.

JUSTIFICATION, VINDICATION AND TYPES OF ETHICAL OUTLOOK

At this point, we can take up the task of reflecting on some different ways in which justificatory and vindicatory stances and their respective ideals may be related to one another. I will sketch out three broad modes of relationship in which they may stand: the priority of the justificatory over the vindicatory; the priority of the vindicatory over the justificatory, and the integration view.

Consider first an ethical outlook in which the justificatory stance and its ideal of righteousness has priority over the vindicatory stance and its ideal of life-affirmation. This can take two general forms. The first is one in which the latter is wholly subordinate to the former such that the character of the vindicatory ideal is itself restructured in terms that make it consonant with the justificatory ideal. This gives rise to a picture of ethics as 'impartial morality' in the sense that critics such as Wolf and Williams take to be problematic. This does not mean that the ideal of life-affirmation does not operate in this ethical outlook at all, rather it means that self-affirmation is pictured in terms of the successful cultivation of one's powers to will and act in ways that are morally justified. We might see the ethical telos of a holy will as one form which this might take and the moral saint (in Wolf's sense) as its embodied expression. This relationship of domination of the justificatory over the vindicatory generates what we may think of as a paradoxical relationship to vindication. I have argued that a natural expression of a course of action being vindicated is the spontaneous positive avowal "I did it!" that expresses first personal pride in the achievement. Notice though that the restructuring of the vindicatory ideal under the domination of the justificatory ideal makes the experience of pride in one's achievement a problem precisely because it would be to express a form of ethical partiality within the domain of moral impartiality in a way that the righteous moral agent cannot coherently do: the moral saint cannot experience pride in being a moral saint because taking up such a prideful relationship to one's own saintliness is incompatible with a saintly outlook in which it is grace that is the relevant relationship.[9]

The second way that the justificatory ideal can stand in a priority relationship to the vindicatory ideal is through a picture in which moral demands have general priority but there is space for allowing agent-centred personal prerogatives. How much space there is, what weight to allow such prerogatives, and indeed whether this can be sustained as a coherent ethical outlook at all are all matters of controversy. However, as Wolf notes, many of those responding to Williams' criticisms of 'impartial morality'

9] This claim may be too expressive of a Christian view of sainthood in which one's saintliness is conceived as a gift of grace rather than a personal achievement.

in its utilitarian and Kantian forms 'have agreed that of course morality should take account of the agent's possible sacrifices, weighing them in the balance against the goals and interests of others that morality is concerned to address and protect.' (Wolf, 2010: 56). Hence this does mark out an effort to articulate a form of ethical outlook that would allow ethical agents to enjoy a form of life-affirmation without generating the vindicatory paradox that marks the domination version of justificatory priority but maintaining the priority of the justificatory stance over the vindicatory stance.

What of an ethical outlook in which the vindicatory stance has priority over the justificatory stance? In its most dominating form, this would lead the character of the ideal of righteousness to be restructured by its subordination to the ideal of life-affirmation. This does not mean that the ideal of righteousness drops out of the picture rather its means that this ideal is reconceived in terms that relate to the agent's cultivation of their powers to realise their valued ends – we might see the ethical *telos* of *kalon* ("noble" or "beautiful" aka "fine") as one form this may take and the Homeric hero as one possible embodied expression of it. Thus, for example, Achilles takes Agamemnon to act unjustifiably towards him but the reasons that this is unjustifiable concern Achilles' personal excellence and contribution to the Greek cause. But less dominating forms are possible. So, for example, Machiavelli's political ethics of glory or Nietzsche's ethics of *amor fati* may be seen as prioritizing the ideal of life-affirmation but acknowledging the ethical demands of what is owed to others. Thus, in *Daybreak*, Nietzsche presents his project of re-evaluation as, in part, oriented to this task: 'we shall restore to men their goodwill towards the actions decried as egoistic and restore to these actions their *value – we shall deprive them of their bad conscience!*' (D s148), while at the same time stressing the following point:

> It goes without saying that I do not deny - unless I am a fool - that many actions called immoral ought to be avoided and resisted, or that many called moral ought to be done and encouraged – but I think that one should be encouraged and the other avoided *for other reasons than hitherto.* (D s103)

The first remark expresses Nietzsche's commitment to the idea of life-affirmation; the second his acknowledgment of the demands of what we owe to each other. [10]

The third position may be thought of as an 'integrationist' view in which the aim is not to prioritize one ideal or the other but to integrate them as far as possible. The two primary exemplars of this view are Aristotle (and, following him, Aquinas) and Hegel (and, following him, Marx).[11] Central to this project is the aim of reconciling the autonomy and flourishing of the individual with the autonomy and flourishing of the community through an understanding of freedom as self-realization where acting

10] We might note, in passing, that the contrast between these first two types of relationship of justificatory and vindicatory stances and ideals denotes the two forms of ethical outlook that Nietzsche considers in his *On the Genealogy of Morality* – slave-morality and noble-morality.

11] I am grateful to Avery Kolers and Cindy Holder for pushing me on this point.

appropriately with respect to others is an integral part of individual self-realization and hence self-realization contributes to the flourishing of the community. Alasdair MacIntyre and Charles Taylor may be seen as contemporary representatives of this approach.

VINDICATION, ETHICS, AND THE MORALITY CRITICS

"Morality", in the sense subject to criticism by a diverse group of philosophers including Alasdair MacIntyre, Charles Taylor, Bernard Williams and Susan Wolf, is understood in the sense of 'impartial morality' (whether in its Kantian, utilitarian or contractualist variants) that I have characterised in the previous section as an ethical outlook that institutes the lexical priority of the justificatory stance to the vindicatory stance. In historical terms, this ethical outlook is a product of European modernity bound up with the emergence of the state and the market as the primary poles of social life (as MacIntyre (2016) stresses) and the demarcation of egoism and altruism as basic action-orientations in social interaction. Rather than mapping this constellation in detail though, what I want to do in this final section is focus on the relationship of the vindicatory stance to two issues that concern all these thinkers, but which are drawn out most specifically by Williams and Wolf, namely, moral luck and meaning-in-life. I focus on these because they help not only to situate my sketch of the vindicatory stance as a dimension of ethical life in relation to the "morality critic" debates but because they also allow me further to clarify the place and character of the concept of vindication in ethical life.

An initial point to make concerns Williams' argument in 'Moral Luck' in which he imagines a fictional version of the artist Gauguin who confronts the fact that he has moral obligations to his family but also has reason to believe that if he stays in France to fulfil those obligations, then he will not be able to try to realise the artistic vocation that he sees as of fundamental value to his sense of his life. It is the case that Gauguin does not, and cannot, know whether he has in within him to be great artist, but he knows that his life will be haunted, perhaps to destruction, by the thought 'if only' if he does not attempt to find out. In Williams' fictional scenario, Gauguin abandons his family, travels to Tahiti to work on his art, and when he eventually returns is recognised as a major artist.

My first concern here is very narrow in that I only want to take up Williams' use of the concept of justification. He writes:

> The justification, if there is to be one, will be essentially retrospective. Gauguin could not do something which is thought to be essential to rationality and to the notion of justification itself, which is that one should be in a position to apply the justifying considerations at the time of choice and in advance of knowing whether one was right (in the sense of it coming out right). (1981, 24)

But does this idea of retrospective justification make sense? Williams' route to it by way of a lengthy detour through the phenomena of regret and agent-regret concludes by specifying the sense of 'justification' in 'retrospective justification' in terms of regret:

> [T]he project in the interests of which the decision is made is one with which the agent is identified in such a way that if it succeeds, his standpoint of assessment will be from a life which then derives an important part of its significance from him from that very fact; if it fails, it can, necessarily, have no such significance in his life. If he succeeds, it cannot be the case that while welcoming the outcome he more basically regrets the decision. [...] That is the sense in which his decision can be justified, for him, by success. (1981, 34-35)

The sense of 'justification' advanced in this remark is not, however, the sense of justification engaged in the earlier passage. It speaks to a situation in which the agent has, all things considered, reason not to regret their action, not to a situation in which the agent confronts the question of whether their conduct meets the requirements of rationally or morally upright conduct. But this is a matter of vindication rather than justification – a point that is illustrated by Williams' acknowledgment of the fact that those (Gauguin's family members) harmed by his decision retain justified grounds for complaint against him.[12] Nor need we, as art lovers grateful that we have Gauguin's art, regard Gauguin's decision as morally or rationally justified; we need merely acknowledge that his actions were vindicated, and that in virtue of the artistic value brought into the world by his decision hold that, all things considered, we have reason not to regret his decision. The problem, I think, with Williams' use of the concept of justification in 'retrospective justification' is that it elides the distinction between justification and vindication – and in doing so suggests misleadingly that 'justification' and 'retrospective justification' are part of the same game. This misleading impression is reinforced when, for example, Williams offers remarks such as the following on questions of moral self-indulgence and of dirty hands:

> One issue that does notably arise with both of these questions, but which, again, I shall not discuss, is the extent to which, and the ways in which, actions offensive to morality can be retrospectively justified – perhaps even morally justified – by success; and what, if they can, may count as success. (1981, 42)

This is ironic since Williams' primary concern was with trying to demonstrate that the claims of 'morality' (understood here in the sense specified by our consideration of 'moral justification') on our lives are not, and cannot be, all that the proponents of impartial morality, and hence of the priority of the justificatory stance, take them to be.

Wolf's reflections on why meaning in life matters helps to indicate in this context that the dimension of human life that pertains to the pursuit of activities of value that

12] While Gauguin has, all things considered, reason not to regret that he has done, it does seem to me he can regret that doing what he has done entailed the moral harm to his family that it did and acknowledge that he owes them recompense. I am grateful to Matt Kramer for pressing me on this point.

we value and that give meaning to our lives picked out by the concept of vindication relates precisely to Williams' key point, namely, that there must be a point at which it becomes "quite unreasonable for a man to give up, in the name of the impartial good ordering of the world of moral agents, something which is a condition of his having any interest in being around in the world at all." (1981, 14) The concept of vindication is consonant with Wolf's argument that meaning 'comes from active engagement in projects of worth, which links us to our world in a positive way' and 'allows us to see our lives as having a point and a value even when we take an external perspective on ourselves' (Wolf 2010, 58). On this view, what gives meaning to our lives thereby gives us reasons to live.

Importantly, however, Wolf also highlights the point that it is unhelpful and obfuscatory to attempt to situate this active engagement in projects of value within the terms of the self-interest/morality framework that "morality" (in the targeted sense) aims to impose. In pursuing projects that give meaning to our lives, we are not pursuing happiness (Wolf, 2010 & 2015: 107-126). An example: when Wittgenstein contracted prostate cancer in 1951, he went to stay with friends in Cambridge and, when he wasn't expected to survive the night and with other friends due to visit the following morning, he made his famous final utterance: "tell them I've had a wonderful life". It has been proposed that this remark must have been intended simply to ease the burden of his friends since much of his life seemed both starkly lonely and intensely unhappy. But the facts of loneliness and unhappiness need not undermine the truthfulness of Wittgenstein's avowal, on the contrary they can serve as markers of the struggle that he underwent to realise what mattered to him – and read in the light of his remark that "Man has to awaken to wonder … Science is a way of sending him to sleep again", the avowal of a 'wonderful' life may have marked a precise and appropriate form of affirmation of the philosophical life that he led.

Wolf puts this point by noting that insofar as pursuit of our projects give meaning to our lives, it is because these projects 'have an independent value that draws us out of ourselves, linking us to larger community or world in a positive way':

> When we act or want to act in the context of these attachments out of love or passion for their objects, we do not do so purely or primarily for our own sakes (not even, therefore, for the sake of being able to live a meaningful life), but at least partly for the sake of the person or project or value that is the object of our love.
>
> If we keep these features in mind, the moralists' injunction that the agent should sacrifice that which gives meaning to his life for the sake of morality is liable to take on a hollow ring. (Wolf 2010, 56-7)

What does Wolf mean here? It is certainly the case that in pursuing our projects – and especially our ground projects (in Williams' sense) – we are doing so because we value the project and understand its value in terms that are not reducible to self-interest, but at the same time the successful realization of (independently) valuable

outcomes through this project is integral to our flourishing and our relationship to it has an ineluctably first personal form expressed, for example, in the spontaneous cry "I did it!". If we read Wolf's use of the phrase 'for the sake of' in instrumental terms, then Wolf's surely correct point is that our relationship to our projects of value does not take the form of aiming at a meaningful life and asking what projects will serve as means to this end. But if we read "for the sake of" as denoting a constitutive relationship between our projects of value and a meaningful life, then acting out of love for, for example, art (Gauguin) or sport (Faldo) just is acting for the sake of a meaningful life. Given her objections to the self-interest/morality framework, I take it that Wolf's concern is to reject the view according to which our pursuit of our projects is pictured in instrumental terms. Gauguin's and Faldo's pursuit of their ground projects is, at once, the pursuit of goals that have independent value and the pursuit of meaningful lives that they can affirm.

This brings me to a second point where engagement with Wolf's work raises a question for what I have characterised as the vindicatory dimension of ethical agency. In sketching this dimension, I have stressed the centrality of 'valuable goals as challenges' and advanced some perhaps overly bold claims about the relationship of vindication and pride and joy in achievement. The question is thus whether 'active engagement in projects of worth' necessarily involves that one is taking up challenging goals?

Wolf's argument would seem to support an affirmative answer to this question. When she introduces the idea of acting out of 'reasons of love' with respect to activities, her examples – for instance, agonizing over the article she is trying to write and struggling to get it right – highlight the point that it is the value of the activity that leads the agent 'to sacrifice ease and exercise discipline in pursuing her goal' (2010). Sacrificing ease and exercising discipline are ways in which we develop our abilities and overcome resistances in pursuing a goal. It is important to notice here that taking up 'goals as challenges' does not mean that one necessarily conceives of the goal as a challenge (although one may) nor that one is engaging in pursuit of the goal because it is challenging and requires the development of one's abilities. Rather one is pursuing the goal because it is valuable.

But two questions arise at this point. First, isn't this talk about challenges all a bit strenuous? What is wrong with wanting just to be? Or to have periods of downtime? Second, what if the agent in question has become a virtuoso practitioner of the activity such that the immediate challenge that her goal poses is no longer difficult for her?

With respect to the first question, the goal of living a simple life is a goal, often quite a challenging one involving considerable self-discipline. To "just be" (in a sense that does not reduce to pure hedonism) is a goal that can be vindicated precisely because it makes demands on one's agency. What, though, of downtime, of taking it easy? We can certainly envisage periods like family summer holidays where we aim simply to let

ourselves relax and recharge[13] – and it may seem odd (unless one finds it really hard to relax or not work) to take pride in this achievement. There are two slightly different points that are relevant here. The first is that the vindication of one's ethical agency is not the only source of life-affirmation. Relaxing and being receptive to, for example, the wonders of nature or of art or of everyday family life can be sources of affirmation that are not ethically vindicatory in character. If we see the desire to 'just be' expressed in terms of downtime and an openness to such sources of life-affirmation, then this the state of 'just being' is a valuable but not ethical vindicatory state. The second is that downtime insofar as it is oriented to recharging, to summoning the renewal of the energies need to pursue one's ground projects, is not vindicated by being successfully realised, rather its vindication is part and parcel of the pursuit of one's valuable goals. It is not an object of freestanding vindication but rather of dependent vindication.

Turning to the second question, we can note that the central issue here concerns the distinction between commitment to a challenging activity and commitment to ensuring that the activity remains challenging, and relatedly the state of boredom.

To get at this question, consider the feelings of frustration and boredom we experience when we are stuck performing an activity that we long ago mastered but are prevented from advancing beyond. Consider, for example, being stuck doing basic mathematics (e.g., multiplication tables) at primary school. However much one loves doing maths, it can be hard to sustain that love in contexts of utter boredom. So what does the bored child (or sensible teacher) do? One thing they might do is to create rules that make it more difficult by, for example, introducing a time-challenge: 'how many questions can I do in under a minute without slips?' This kind of response is a maths-loving response where the love is expressed by trying to make the activity more challenging by requiring a heightened level of attention. The bored child is acting in a way that is designed to sustain their love of the activity. This kind of example suggests that love, meaning, and challenge are related such that if an activity one values ceases to be challenging then it could cease to be a source of meaning but also that loving an activity will be expressed by attempting to sustain its challenging quality.

But we can also think of examples that concern challenging activities where the activity involves a range of skills and demands a level of attention – e.g., writing a philosophy article or making a complicated dessert – such that one can experience pride and joy is just exercising and sustaining the skills and attentiveness required to perform the activity well. It is, we might say, an intrinsically challenging activity (so we may value the ideal of effortless performance in relation to such activities precisely because of the challenge posed by such an ideal and the demands that it makes on us).

13] Although as Matthew Kramer rightly pointed out to me, "Downtime" spent with one's spouse and/or children is not altogether uninvolved in the meeting of challenges. Rather, it is partly constitutive of one's efforts to meet the challenge of sustaining a robust and loving marriage and/or the challenge of sustaining a close-knit family.

In these kinds of cases, boredom does not denote that the activity has ceased to be challenging, rather it discloses that what one values has changed.

From these reflections, we may say that if it is reflection on Williams' work on moral luck that helps make clear the place of vindication in ethical life; it is Wolf's reflections on meaning-in-life that deepen our appreciation of its ethical character.

CONCLUSION

The argument I have offered here has aimed to clarify and characterise the place of vindication in ethical life. I began by trying to sketch out the contours of the concept of vindication in order to be able to distinguish the reflective stance that this concept institutes from that constructed by the concept of justification. I have then proposed that whereas moral justification involves a stance towards one's ethical agency oriented by the ideal of upright conduct, vindication involves a relationship to one's agency oriented by the ideal of life-affirmation. This led me to the claim that these two stances and their respective ideals can be placed in more than one kind of relationship to each other – and I sketch out three general forms that this relationship may take. Finally, I situated this focus on vindication in relation to some of the concerns of the "morality critics", particularly Bernard Williams on moral luck and Susan Wolf on meaning-in-life. There is no doubt much more that could be said, and it must be acknowledged that this argument still has a rather provisional character, but I hope that it provides a basis for further investigation of the salience and significance of the concept of vindication for ethical life.

dowen@soton.ac.uk

REFERENCES

Adams, R.M. 2010. Comment. In S. Wolf *Meaning in Life and Why It Matters*, Princeton, Princeton University Press, 75-84.

MacIntyre, A. 2016. *Ethics in the Conflicts of Modernity*. Cambridge, Cambridge University Press.

Owen, D. 2009. Review of 'Reginster. B. (2006) *The Affirmation of Life: Nietzsche on Overcoming Nihilism*, Cambridge MA., Harvard University Press.' *Ethics* 119(3): 598-602.

Reginster. B. 2006. *The Affirmation of Life: Nietzsche on Overcoming Nihilism*, Cambridge MA. Harvard University Press.

Strawson, P. 1974. Social Morality and Individual Ideal. In *Freedom and Resentment and other essays*, London, Routledge, 29-49.

Williams, B. 1981. *Moral Luck*, Cambridge, Cambridge University Press.

Wolf, S. 2010. Meaning in Life and Why It Matters. In *Meaning in Life and Why It Matters*, Princeton, Princeton University Press, 1-63.

———. 2015. Happiness and Meaning: Two aspects of the good life. In *The Variety of Values*, Oxford, Oxford University Press, 107-26.

Public Reason 14(2)-15(1): 29-34

The Self, Love and the Other: Thoughts on Nietzsche, Kant and Owen

Joe Saunders
Durham University

Abstract. Owen (2009, 2017) contrasts Kant and Nietzsche's strategies for dealing with self-love. Kant sees our self-love as ineliminable, and looks to adopt a strategy of subordinating or supressing it to the moral law. Owen sees Nietzsche, by contrast, as adopting a strategy of *channelling* self-love, directing it in ways that serve both individual and collective development. In this paper, I argue that we do not need to just supress, subordinate, or channel self-love, for we can also move away from it. Doing this involves recognising our own dependence on and vulnerability to each other, and empathy. I contend that emphasis on these things also moves us away from Kant and Nietzsche.

Keywords: Nietzsche, Kant, self-love, empathy.

Human beings can be self-absorbed, selfish, and prone to self-love. How are to deal this feature of our lives? Owen (2009, 2017) contrasts two strategies for dealing with self-love. Kant sees our self-love as ineliminable, and looks to adopt a strategy of subordinating or supressing it to the moral law. Nietzsche, by contrast, adopts a strategy of *channelling* self-love, directing it in ways that serve both individual and collective development. In this paper, I make the case for an alternative option. I argue that self-love is not ineliminable in the way that these thinkers hold it to be. And thus we do not need to just supress, subordinate, or channel it – we can also move away from it. Doing this involves recognising our own dependence on and vulnerability to each other, and empathy. I contend that emphasis on these things also moves us away from Kant and Nietzsche.

The paper takes the following structure. I begin by briefly outlining Kant's, Nietzsche's and Owen's views on the topic (§1). I then argue that self-love is not as ubiquitous as they make out (§2), before turning to look at empathy, vulnerability and dependence (§3). Finally, I contend that we should move towards emphasising these things, and away from Kant and Nietzsche's emphases on independence, autonomy, self-mastery, and self-responsibility (§4).

I. KANT, NIETZSCHE AND CHANNELLING SELF-LOVE

Kant maintains that we either act from self-love, or duty.[1] These are the two options for human beings. We possess the capacity to act from duty. But when we go wrong, and fail to act from duty, we act from self-love. Kant famously thinks that acting from duty

1] See, for instance, Kant's claims at V: 22-23 in the *Critique of Practical Reason*; for further discussion, see Sticker and Saunders 2022.

alone has moral worth, and so our self-love ought to be suppressed or subordinated to the moral law.[2]

Instead of looking to supress or subordinate self-love, Nietzsche hopes to channel it. And Owen (2009, 2017) provides a compelling account of how this occurs, and how channelling self-love can serve both individual and collective development.

One important part of this involves self-respect:

> A preliminary view concerning the relationship between self-love and self-respect would thus be that valuing the disposition of *amor fati* [love of fate3]broadly equates to valuing the will to self-responsibility (i.e., ethical autonomy) and, hence, that self-love consists in valuing self-respect. (Owen 2009, 216)

This seems plausible. Self-respect is healthy, especially in contrast to pernicious forms of self-hatred and self-denial. And insofar as we can channel self-love towards self-respect, that seems like a good strategy.

What about the other parts of Nietzsche's view, the emphasis on self-responsibility and autonomy? It is worth noting that his conception of autonomy differs from Kant's:

> […] in contrast to Kant, [Nietzsche] takes the achievement of autonomy to be the ongoing achievement of standing in a relationship of self-responsibility and self-overcoming to oneself such that the practice of certain virtues are integral to developing and sustaining this relationship – and that these virtues include those cultivated by an agonal culture such as courage and independence of mind, that is, those virtues integral to taking up and overcoming challenges. (Owen 2017, 153)

Owen draws upon Nietzsche's' interest in ancient Athens *agonal culture* to suggest how self-love can be channelled. An agon is a site of competition, contest, or conflict, and Owen notes that, for Nietzsche:

> An agonal culture […] cultivates just that will to self-overcoming which is the disposition of freedom. Nietzsche takes certain passions – desires for respect, honour, glory – to be channelled by the agon in ways that serve culture, society and polity both by cultivating the appropriate practical relation to self in participants and by developing the excellence of practices (art, politics, etc.). (Owen 2017, 149)

More precisely, the agon:

> […] acts to cultivate will to power (the instinct for freedom) in such a way that the feeling of power (the feeling of effective agency) tracks power (effective agency) *and*, at the same time, supports a practical relationship to self in which the power to engage in the self-directed development of one's powers (and hence the dispositions of self-responsibility and of self-overcoming) is central to one's ethical outlook. (Owen 2017, 149)

2] How exactly this works is complicated. One of Kant' suggestions is that the moral law strikes down our self-conceit, which humiliates us, and produces a special feeling of respect for the moral law, which we then go to act on; see V: 73 in the *Critique of Practical Reason.*

3] For a thoughtful discussion of Amor Fate, see Hans-Pile 2011.

So conceived, our desires for respect, honour and glory, and our will to power can be helpful. Here we see that, for Nietzsche self-love can be channelled towards self-mastery and self-responsibility, in a way that benefits the individual, and can even serve culture and society more broadly. We see an ethical outlook in which self-love isn't necessarily problematic, and where self-mastery and self-responsibility are central.

I want to suggest an alternative ethical outlook. But before I do that, let's stop to think about just how ubiquitous self-love really is.

II. IS SELF-LOVE UBIQUITOUS?

The view that self-love is ubiquitous is common in the Lutheran tradition. We find traces of this view in Kant, Nietzsche, and Murdoch, amongst others.[4] I want to put some simple pressure on this view. For I don't think self-love is as ubiquitous as these thinkers make out. I don't deny that it exits, but think the evidence suggests it is less widespread than we might worry. I also contend that where it does exist, we can look to overcome it.

The view that we are always motivated by self-love (or self-interest) is a form of psychological egoism. And this is typically seen as an under-motivated view, in its own way, a sort of scepticism about our motivations. After all, perhaps we are all only motivated by self-interest. That's a possibility, but why is it a possibility that we ought to take seriously?

Here it will help to interrogate the nature of the claim a little bit. What kind of the claim is the claim that we are always motivated by self-love? One option is that it is a sort of conceptual claim: when one acts, one does what one prefers to do, and this is a form of self-love. So conceived, we always act from self-love, but this doesn't tell us anything substantial about our actions or motivations.

The other option is that it is a substantial, perhaps empirical or psychological, claim. Other beings might be able to be motivated by other concerns, but us human beings, as a matter of fact, are motivated by self-love. This is a more informative claim, that tells us something substantial about our motivation, but one that we would need evidence for. And, as far as I can tell, the evidence doesn't seem to back it up.

Of course, one might object that psychological egoism is not falsifiable in this sort of way. If that's true, it takes us back to viewing it as a conceptual, perhaps a priori, claim about the nature of our motivation. But that cuts both ways; it avoids being falsifiable, but in doing so, it also lacks any empirical support. I think we should move away from such an approach, opening up our claims about ourselves to the world and evidence.

The evidence suggests that we can often be self-interested, but are not always. For human beings can be motivated by the thoughts and feelings of others, which we'll look

4] For a helpful account of Luther's influence in German Philosophy, see Robert Stern, "Luther's Influence on Philosophy" *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.).

at in more detail when we turn to empathy in the next section of the paper. Miller (2018, 109) offers one example, where he considers the motivations behind lying. According to one study (De Paulo et al. 1996, 987), 57% of the time we lie, we do so for self-orientated reasons, but 24% of the time, we do so for others, looking to protect their feelings for instance. Now it is not clear whether this lying is right or wrong, for that we would need to know the full context, but either way it does suggest that we are not always motivated by self-love. Empathy suggests that we can be well-motivated by other things than self-love. And the case of lying to protect others feelings also suggests that we can be poorly motivated by more than just self-love as well. I contend that we go wrong and right in a variety of ways, and not just due to self-love.[5]

III. HOW DO WE MOVE AWAY FROM SELF-LOVE?

So far, I have argued that self-love is not as ubiquitous as we might worry. But I have not argued that it doesn't exist. It does. It is fairly widespread, and can be a serious problem. Thankfully though, we can look to overcome it. How? Well, at the most basic level, by focussing on something outside of the self. Simone Weil offers an influential account of this in her 'Reflections on the Right Use of School Studies with a View to the Love of God'. She champions the importance of attention, which she claims destroys the evil in us. For Weil, attention takes us out of ourselves:

> Attention consists of suspending our thought, leaving it detached, empty, and ready to be penetrated by the object. (Weil 1951, 111)

> Above all our thought should be empty, waiting, not seeking anything, but ready to receive in its naked truth the object that is to penetrate it. (Weil 1951, 112)

For Weil, this is true of attention in general. It is good for us to really pay attention to a problem in maths, as it takes us out of ourselves. The general point about de-selfing is taken up by Murdoch in a famous passage in *The Sovereignty of the Good*:[6]

> I am looking out of my window in an anxious and resentful state of mind, oblivious of my surroundings, brooding perhaps on some damage done to my prestige. Then suddenly I observe a hovering kestrel. In a moment everything is altered. The brooding self with its hurt vanity has disappeared. There is nothing now but kestrel. And when I return to thinking of the other matter it seems less important. And of course this is something which we may also do deliberately: give attention to nature in order to clear our minds of selfish care. (Murdoch 1970, 82)

For Murdoch, it is paying attention to nature that takes us out of ourselves. This strikes me as a little more promising than maths.

5] For an extended discussion of this, with particular reference to Kant's claim that we always go wrong through self-love, see Sticker and Saunders 2022.

6] Murdoch is an interesting case, because she is also quite Lutheran is her assessment of human nature, and how egotistic we are. For further discussion, see Stern 2022.

However, there is a simpler answer. We get out of our selves and our own concerns, by opening ourselves up to other people. And contemporary moral psychology seems to back this up: through affective empathy, we are open to being moved by others. Here is Elisa Aaltola on this point:

> […] affective empathy is intrinsically involved and other-directed; moreover, it is redolent of openness toward others, consists of it. This is quite simply because, in a very tangible fashion, affective empathy opens us to the influence of others by causing us to resonate with their emotive states. Thereby it impels one to become exposed or receptive to the other, i.e. to allow the other to bear an impact on oneself (hence resisting detachment), and to note and pay heed to others' experiences (hence making other-directedness possible)." (Elisa Aaltola 2014, 247)

This is good news: we are not just self-absorbed creatures. We are capable of being deeply receptive to the thoughts and feelings of others.[7]

This is related to recognising our own dependence and vulnerability. And here, we see an emphasis on aspects of us that depart from Kant and Nietzsche's emphasis on independence, autonomy, and self-mastery. It's worth thinking a little bit more about this. After all, what is so valuable about our vulnerability and dependence. Aren't those just facts about us? And perhaps facts that we should strive to overcome.

No. For one, as Aaltola suggests, they make us better people. How? How does recognising our vulnerability and dependence make us better people? Here's Monique Wonderly on this point:

> Recognising vulnerability and dependence as central features of our own lives allows us to see others' vulnerabilities as evidence of a shared condition between us. Taking up this shared condition is thought to be key to motivating caring attitudes and behaviors towards those who require aid (Wonderly 2022, 993)

This resonates with what we saw earlier from Aaltola:

> […] affective empathy opens us to the influence of others by causing us to resonate with their emotive states. Thereby it impels one to become exposed or receptive to the other, i.e. to allow the other to bear an impact on oneself (Elisa Aaltola 2014, 247)

This is what contemporary moral psychology suggests. And it makes sense. We live together, and are affected by each other. We are dependent vulnerable creatures, exposed and receptive to each other. Acknowledging this is good for us, both individually and collectively.[8]

7] Of course, there are exceptions. For a comprehensive discussion of moral responsibility and psychopathy, see Baxter (2022). And for an intriguing defence of an agent without such emotional capacity, see Sofronieva (2021). Sofronieva's defence centres around the admirable sociopathic character Amos from the tv show *The Expanse*. But also, see Singh 2021, who argues that Amos has failed to properly channel his will to power, and thus falls short as a Nietzschean ideal.

8] This can also lead to the formation of a *we*, which might be thought to expand self-love in a non-pernicious way. Cf. Walther 1922 for an account of ethical community. And see Saunders 2024 for discussion of this issue, including the worry that such a community or we might unhelpfully engulf the self.

Alongside this important moral benefit to recognising our vulnerability and dependence, I think there's something touching to our condition. Here is an uncharacteristic passage from Kant, from the *Metaphysics of Morals*:

> Would it not be better for the well-being of the world generally if human morality were limited to duties of right, fulfilled with the utmost conscientiousness, and benevolence were considered morally indifferent? It is not so easy to see what effect this would have on human happiness. But at least a great moral adornment, benevolence, would then be missing from the world. This is, accordingly, required by itself, in order to present the world as a beautiful moral whole in its full perfection, even if no account is taken of advantages (of happiness). (VI: 458. 2-11)

It's uncharacteristic, given Kant's emphasis on (perfect) duties, and the strict priority these have over all other elements of our lives. But nevertheless, even Kant recognises that without benevolence, "a great moral adornment" would be missing from the world. I think something similar is true of our dependence and vulnerability. One can imagine a world without these things, where we are more independent and less vulnerable. And there would be benefits to this. For life hurts, and it would be no insignificant gain to insulate ourselves from these hurts.[9] But we'd lose something in the process, something touching about our state, the type of creatures we are, creatures who need one another.

IV. TIME FOR ANOTHER RE-EVALUATION OF VALUES?

In Kant, I see an emphasis on reason, freedom and autonomy. In Nietzsche, Owen sees an emphasis on autonomy, independence, self-responsibility and self-mastery. These are good things. And at certain times, for certain people, it would be beneficial to emphasise these things, and encourage people to be more autonomous, independent and to develop self-mastery.[10] I worry that now is not that time.

We live in an increasingly individualistic age.[11] We are increasingly isolated, and encouraged to fend for ourselves. In this age, I think we need an emphasis on different things. I think we need more emphasis (especially amongst men) on our vulnerability, dependence, and empathy.

Johnathon Wolff (2015) argues that one of, if not *the* most important things that political philosophers can do, is to emphasise a value that has been relatively neglected. He draws upon an insight from Margaret MacDonald:

> The value of the political theorists, however, is not in the general information they give about the basis of political obligation but in their skill in emphasizing

9] We'd also safeguard ourselves from some of the downsides of empathy; see Cuff et al. (2016, 149) for discussion of some of these downsides.

10] As we say in §1, Owen (2009, 216) points out that part of self-love is self-respect, and that also seems very important for people to emphasises and develop.

11] See, for instance, Judt 2010.

> at a critical moment a criterion which is tending to be overlooked or denied. (MacDonald 1940, 112)

Wolff notes that the latter half of the twentieth century emphasised the value of individual responsibility, which ended up doing considerable damage to the welfare state, and large numbers of people who live in poverty. He does not deny that individual responsibility is a valuable thing, but he thinks its value has been overemphasised, to the detriment of other values, such as compassion and solidarity, and now is the time to emphasise these.

I agree with Wolff, both in his methodological point about what philosophy can do, and in the substantial point about the importance of emphasising compassion and solidarity now. And I think these points both apply roughly in this paper. I worry that Nietzsche (and Kant) overemphasise the importance of individual responsibility, amongst other things.

I worry that Nietzsche lionises strength, pride, and independence. I worry that he champions self-mastery and self-responsibility to the detriment of other important features of our lives: our weakness, our vulnerability, and our dependence on each other.[12]

But am I being unfair to Nietzsche here? After all, he does not only lionise these things. At times, he writes beautifully of a fuller picture of human life:[13]

> He who knows how to regard the history of man in its entirety as his own history, feels in the immense generalisation all the grief of the invalid who thinks of health, of the old man who thinks of the dream of his youth, of the lover who is robbed of his beloved, of the martyr whose ideal is destroyed, of the hero on the evening of the indecisive battle which has brought him wounds and the loss of a friend. […] to take all this upon his soul, the oldest, the newest, the losses, hopes, conquests, and victories of mankind: to have all this at last in one soul, and to comprise it in one feeling: - this would necessarily furnish a happiness which man has not hitherto known, - a God's happiness, full of power and love, full of tears and laughter, a happiness which, like the sun in the evening, continually gives of its inexhaustible riches and empties into the sea, - and like the sun, too, feels itself richest when even the poorest fisherman rows with golden oars! This divine feeling might then be called humanity! (GS §337)

And while he is very critical of (what we would now call) sympathy, he does have some intriguing things to say about (what we would now call) empathy.[14] Moreover, even though he is critical of sympathy, pity and compassion, he does express some compassionate thoughts:

12] For an overview of an alternative ethics, one that focuses upon situations where we are powerless, see Batho 2015.

13] See also the learning to love passage (GS §334); thanks to David Owen for pointing me towards these passages.

14] See D §142, and Ganesh 2017 and Özen 2021 for discussion.

> What dost thou think most humane? To spare a person shame. (GS §337)

There is more to be said about weakness in Nietzsche as well.[15] At one point, Owen (2009, 216) draws upon passage 290 from the *Gay Science* to suggest that:

> Nietzsche distinguishes between non-necessary weaknesses that can be removed or overcome and necessary weaknesses that are to be concealed or reinterpreted and made sublime. (Owen 2009, 216)

Owen (2009, 217) then continues to approvingly cite Ridley (2005) who claims:

> What is not necessary, and is weak or ugly, should be removed. What is necessary should, if weak or ugly, either be concealed […] or else 'reinterpreted', so that one learns how to see it as beautiful, as a strength.

This seems to allow for some valuing of weakness.[16] But both Nietzsche and Ridley suggest that if weakness is not necessary, it should be overcome. And when it's necessary and can't be overcome, it should be concealed or reinterpreted as strength.

On the first point, I'm not sure that our vulnerability and dependence is strictly necessary. Perhaps we could strive to overcome it, and become more independent and insulated from each other. For some people, in some situations, that will be helpful. But I think it can also be a mistake. For, as we have seen, this is an important part of who we are, it brings us together, and connects us.

As for the second point, perhaps Nietzsche, Owen, Ridley and I end up in a similar place here, where we can all agree that our vulnerability and dependence are good things. For Nietzsche, it might be a weakness, but one that we can reinterpret as a strength.

If that's right, then the key difference is one of emphasis. Nietzsche provided some helpful correctives to the self-denial and self-hatred he found in certain forms of Christianity, with his emphasis on self-love, pride, honour, self-responsibility and self-mastery. But maybe it's time to emphasise other values now. Allow me end with such an emphasis, a beautiful passage by John Jeremiah Sullivan, talking about Jesus:

> "The latchet of whose shoes I am not worthy to unloose." I can barely write that. He was the most beautiful dude. […] His breakthrough was the aestheticization of weakness. Not in what conquers, not in glory, but in what's fragile and what suffers—there lies sanity. And salvation. "Let anyone who has power renounce it," he said. "Your father is compassionate to all, as you should be." (Sullivan 2012, 33)

joe.saunders@durham.ac.uk

15] See Owen 2005 for further discussion of weakness in Nietzsche.

16] Of course, this is not the only passage in Nietzsche on weakness. At one point, Owen discusses another passage from the *Genealogy of Morals*, which seems more dismissive of weakness: "[…] it enables the majority of mortals, the weak and the down-trodden of all sorts, to practise that sublime self-deception – the interpretation of weakness itself as freedom, of the way they simply are, as merit." (GM1 §13); for discussion, see Owen 2014, 78-79. See also GS §357.

REFERENCES

Aaltola, Elisa. 2014. Varieties of Empathy and Moral Agency. *Topio* 33, 243-53.

Batho, David. *The phenomenology of powerlessness*. Colchester: The University of Essex, 2015.

Baxter, Jim. *Moral Responsibility and the Psychopath: The Value of Others.* Cambridge: Cambridge University Press, 2022.

Cuff, Benjamin MP, Brown, Sarah J, Taylor, Laura, and Howat, Douglas J. 2016. Empathy: A review of the concept. Emotion review 8(2): 144-53.

Ganesh, Akshay. 2017. Nietzsche on Honor and Empathy." *Journal of Nietzsche Studies* 48(2): 219-44.

Han-Pile, B. 2011. Nietzsche and amor fati. *European Journal of Philosophy*, 19(2): 224-61.

Judt, Tony. 2010. *Ill fares the land*. Penguin.

Kant, Immanuel. 1996. *Critique of Practical Reason*, translated by Mary Gregor. Cambridge: Cambridge University Press.

———. 2019[1785]. *Groundwork for the Metaphysics of Morals*, translated by Christopher Bennett, Joe Saunders and Robert Stern. Oxford: Oxford University Press.

———. 1996[1797]. *Metaphysics of Morals*, translated by Mary Gregor. Cambridge: Cambridge University Press.

MacDonald, Margaret. 1940. The language of political theory. In *Proceedings of the Aristotelian Society*, vol. 41, 91-112.

Murdoch, Iris. 1970. *The Sovereignty of the Good*. London: Routledge.

Nietzsche, Friedrich. 1997. [D] *Daybreak: Thoughts on the prejudices of morality*. Cambridge University Press.

———. 1974. [GS] *The gay science: With a prelude in rhymes and an appendix of songs*. Vol. 985. Vintage.

———. 1989. [GM] *On the genealogy of morals and ecce homo*. Vintage.

Owen, David. 2005. On failing to be agents: freedom, servitude, and the concept of "the Weak" in Nietzsche's practical philosophy. *Philosophical Topics*, 33(2): 139-59.

———. 2009. Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency. *Nietzsche on Freedom and Autonomy,* edited by Ked Gemes and Simon May, 197-222. Oxford: Oxford University Press.

———. 2014. *Nietzsche's Genealogy of Morality*. London: Routledge.

———. 2017. Freedom as Independence. In *Nietzsche's Engagements with Kant and the Kantian Legacy*, edited by Marco Brusotti, Herman Siemens, João Constâncio, Tom Bailey, Maria João Mayer Branco, and Katia Hay, 137-158. London: Bloomsbury.

Özen, Vasfi O. 2021. Nietzsche's Theory of Empathy. *Philosophical Papers* 50(1-2): 235-80.

Ridley, Aaron. 2005. Introduction. In *Nietzsche, Friedrich Wilhelm, The Anti-Christ, Ecce Homo, Twilight of the Idols, and Other Writings*, edited by Aaron Ridley, and Judith Norman. Cambridge: Cambridge University Press.

Saunders, Joe. 2024. Ethics. In *Oxford Handbook of Nineteenth Century Women Philosophers in the German Tradition*, edited by Nassar, D. and Gjesdal K. Oxford: Oxford University Press.

Singh, Pankaj. 2021. Amos Meets Nietzsche. *The Expanse and Philosophy: So Far Out into the Darkness,* edited by Jeffrey L. Nicholas, 57-64. Wiley Blackwell.

Sofronieva, Diana. 2021. Is Amos Evil?. *The Expanse and Philosophy: So Far Out into the Darkness*, edited by Jeffrey L. Nicholas, 65-73. Wiley Blackwell.

Stern, Robert. 2020. Luther's Influence on Philosophy. *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), edited by Edward N. Zalta, URL = <https://plato.stanford.edu/archives/fall2020/entries/luther-influence/>.

———. 2022. Murdoch and K. E. Løgstrup. In *The Murdochian Mind*, edited by Panizza, S.C. and Hopwood, 331-344. M. Routledge, Abingdon.

Sticker, Martin., and Saunders, Joe. 2022. Why we go wrong: beyond Kant's dichotomy between duty and self-love. *Inquiry* 68(2): 794-825.

Sullivan, John Jeremiah. 2012. *Pulphead: Notes from the Other Side of Americ*a. Random House.

Walther, Gerda. 2021[1922]. A Contribution to the Ontology of Social Communities. Translated by Anna Ezekiel in *Women Philosophers in the Long Nineteenth Century: The German Tradition*, edited by Dalia Nassar and Gjesdal Kristin. New York: Oxford University Press.

Weil, Simone. 1951. *Waiting on God*. Translated by Leaslie A. Fielder. New York: Harper and Row.

Wolff, Jonathan. 2015. Political philosophy and the real world of the welfare state. *Journal of Applied Philosophy* 32(4): 360-72.

Wonderly, Monique. 2022. Love and the Anatomy of Needing Another. In *Oxford Handbook of Moral Psychology*, edited by John Doris and Manuel Vargas, 983-99. Oxford: Oxford University Press.

Public Reason 14(2)-15(1): 19-24

Does Empathy Offer a Good Alternative to the Concept of Self-Love in Kant and Nietzsche's Views? A Reply to Saunders

Allan Arturo Gonzalez Estrada
Universidad Nacional, Costa Rica

Abstract. Recent discussions on the views of self-love in Kant and Nietzsche (Owen 2017, 2019; Saunders, 2023) suggest a conflict between this concept and our duties to others and, how we should pursue and implement such responsibilities at individual and social level. If Kant is correct, our self-love should be subordinated to the moral law or practical reason; otherwise, self-love could influence the will and compromise free choice. On the other hand, as Owen has pointed out, Nietzsche offers an alternative approach to channelling self-love. In this Nietzschean view, self-love becomes the cornerstone of our moral and social development. Saunders argues that empathy, as a third alternative, can help us escape this dichotomy. However, I contend that empathy, while important, is not sufficient, as it may be a fleeting emotion that fails to convey dignity and respect to others. I will argue that a Nietzschean perspective on self-love, combined with Kant's doctrine of virtues, helps us understand the current debate. My view is that if a bridge, between the Duties of Love and the Duties of Virtue toward other human beings in a Kantian framework with the channelling strategy of Nietzsche can be created, then we can keep the concepts of dignity and respect and over empathy. Once this bridge is established, the Duties of Virtue toward other human beings can effectively be utilized to pursue personal and collective development within the Nietzschean position, this approach aligns with a political and ethical view aimed at elevating morality via dignity and respect to the highest possible virtue for all.

Keywords: moral, empathy, universal law, self-love, duties, respect, dignity, Kant, Nietzsche, virtue, love, happiness.

In recent days (October 2023), the border between Colombia and Panama, particularly the region known as the Darien Gap, has been in the centre of the news as thousands of migrants – fathers, mothers, children, and whole families – make a hazardous trek through the Darien Jungle trying to reach the United States of America. Such a barbaric picture could suggest a total breakdown of the principles of humanity, such as religious, political, cultural, social, and economics, that are supposed to be at the centre of the inception of our modern societies – or at least, should be. In such a state of total helplessness, perhaps our first reaction is a feeling of empathy. But it is also true that when we see people in such a state of despair, we could think about the inequalities, the political, economic, and relational dominations that ignite such a massive exodus of people. Hence, may be worth review the ideas of Nietzsche regarding autonomy and self-love, to find a path to elevate our individual and collective moral development. However, even if this is a departing point, this is not enough. It Could be necessary to explore a path where feelings like empathy could serve not only as a motivation to discover the true value of others, but also to empower them, as Kant suggests could be done via dignity and respect (duties of virtue toward other human beings) and promote real change in the lives of people. This reply to Joe Saunders and

its views on empathy will explore (I) the standpoint of self-love in Kant, Nietzsche, and the interpretation of Owen. In (II), we will discuss the role of empathy, as presented by Joe Saunders, but I suggest only as the starting engine of the duties of virtue toward other human beings, in (III) I will outline the idea that dignity and respect is a bridge that work as empowerment tools to channel self-love.

I. SELF-LOVE AND COMMUNITY

The central concept in Aristotle's "Politics" is that of community (κοινωνία), the notion that human beings come together in social organizations with a vision of some common good (1252a) as the goal of such social structures. However, as humans, these communities are formed by individuals with emotions, feelings, and diverse interests and predispositions[1], how can these differences be set aside to pursue this common good? In his work "Idea for a Universal History with a Cosmopolitan Aim," Kant introduces the idea of "unsocial sociability" (ungesellige Geselligkeit) as a fundamental characteristic of human beings (Owen, 2017). In other words, as Kant suggests, this concept of unsociability includes a "trait that predisposes them to want to direct everything only to their own ends." (8: 21). However, if this predisposition to pursue one's own ends is fundamentally rooted in self-love, and if self-love can potentially compromise our will and hinder our ability to choose freely in alignment with universal moral laws, then the only viable option according to Kant, as Owen proposes, "is to find a way to integrate a form of self-love that is subject to the guidance of practical reason" (Owen, 2017). However, it is also worth considering that the concept of self-love can be examined from a more constructive perspective. In other words, even if practical reason can regulate the predisposition of self-love, the idea of unsocial sociability implies more complex relationships that involve inequality and domination, demanding a deeper analysis. I may agree with Owen, who claims that Kant may not be well-suited to address the broader forms of domination, which Nietzsche highlights concerning economic and social forms of radical dependency (Owen, 2017) that can arise in our community life. Nonetheless, as I am attempting to argue, this Nietzschean account could be useful as we recognize the value of the duties of virtue in general terms, as outlined by Kant, which can serve as a bridge to connect with Nietzsche's more empowering ideas.

1] In accordance with Kant's terminology, the difference between predisposition (Anlage) and propension (Hang) is significant. Kant defines predispositions as follows: "By the predispositions of a being, we understand not only its constituent elements, which are necessary to it, but also the forms of their combination, by which the being is what it is. They are original if they are necessarily involved in the possibility of such a being, but contingent if it is possible for the being to exist without them." In contrast, Kant defines propension by stating: "By propensity (propensio), I understand the subjective ground of the possibility of an inclination" (RGV 46-47).

This distinction holds importance because if we consider empathy as a form of propension legislated in terms of moral law, then according to Kant, it is "regarded as having been acquired (if it is good)" (RGV 47-48). It is worth noting that empathy is indeed considered a virtue.

Empathy can form the foundation but should not be considered the sole structural support for the concept of community and our social relations.

II. JOE SAUNDERS' VIEWS ON EMPATHY

Joe Saunders' motivation (2023) for presenting an alternative perspective to Kant and Nietzsche on self-love stems from the idea that self-love is not inherently good or bad in human beings. Instead, the actions we take are contextually grounded. Thus, self-love cannot be considered the sole or primary motivator for our decisions. As Saunders suggests, we may sometimes act in seemingly selfish ways to help and protect others. Therefore, according to Saunders, empathy, as he correctly points out, can motivate us to act based on our consideration of the thoughts and feelings of others (Saunders, 2023).

Saunders' idea revolves around recognizing our interdependence and vulnerabilities in others, which is undoubtedly a noble perspective. However, empathy it may not be sufficient to address and rectify the asymmetrical relations that emerge within communities, since empathy is a 'local attitude,' its focus of attention is consequently limited. To illustrate this point, let's revisit the example I used earlier – the people crossing the Darien jungle. It is unquestionable that their actions result from forms of radical dependency arising from social, economic, and cultural factors in their home countries. In the first place, one could suggest that these people act out of self-love as they attempt to protect themselves and their loved ones. However, in doing so, they expose themselves to the dangers and challenges of crossing a jungle. Our empathy for them, as a feeling that arises when we witness such news, is a natural response. But as Kant suggests in the Metaphysics of Morals:

> In fact, when another suffers, and though I cannot help him, I let myself be affected by his pain (through my imagination), then two of us suffer, even though the trouble only affects one. But there cannot possibly be a duty to increase the ills in the world and do good out of compassion (6: 457).

In other words, empathy is a feeling that should underpin our social interactions, but it may not directly solve the problems faced by the person who is suffering, they need some sort of empowerment[2]. As Nietzsche suggests, we need to empower ourselves through the creative force of ressentiment, a natural force (Owen, 2017). This concept of empowerment is crucial to understanding how empathy may or may not work, but most importantly to work in relation to others, otherwise, we could end with a view of Adam Smith:

2] It is necessary to clarify the concept of 'empowerment' beyond common speech. In the context of forced migrants, empathy alone proves insufficient; rather, a transfer of power is essential. This transfer allows them to overcome, even partially, the conditions of impoverishment that force them to leave their countries. From a political perspective, 'empowering' entails redistributing resources and promoting capabilities.

> Man is assigned a role that is much humbler but also much more suitable to the limited nature of his powers and his intellect – namely the care of his own happiness and of the happiness of his family, his friends, his country. (Smith, p 215)

So, empathy could have a limited scope in terms of people and time; therefore, it is necessary to explore the functional basis of empathy as a foundation to empower people.

III. THE STRUCTURAL BASE OF EMPATHY TO EMPOWER THE OTHERS WITH DIGNITY AND RESPECT AS A BRIDGE BETWEEN KANT AND NIETZSCHE

While looking for an alternative to move away from self-love, as Saunders suggests, can be a compelling argument, it is also important not to multiply the causes or beings. Empathy is undoubtedly at the foundation of our social relations. As human beings, we are, as Saunders has explained, motivated by the thoughts and feelings of others[3]. However, there are times when empathy alone may not suffice to help them. In such situations, it might be more beneficial to organize other members of the community. For example, historically, charity has been considered a form of justice, as was common in the 19th century. But even this did not seem to provide a complete solution to the problem at its core. So, what philosophical tools can we employ to chart a better course?

Kant presents the concept of the duties of virtue towards others. When we help others, they retain "their respect for themselves" (6: 449). It seems to me that this is the key point. A person may act in a selfish way when they lose their dignity and respect, feeling abandoned by the members of the community[4]. Therefore, duties of dignity and respect towards other human beings serve as tools to empower those who are suffering or distressed, because other feelings may not work, in The *Metaphysics of Moral*, Kant suggests that,

> Humanity itself is a dignity; for a human being cannot be used merely as a means by any human being (either by others or even by himself) but must always be used at the same time as an end... Hence, there rests on him a duty regarding the respect that must be shown to every other human being (6: 463).

3] Note for example the idea of empathy sketched by Marcus Aurelius: "Acquire the habit of attending carefully to what is being said by another, and of entering, so far as possible, into the mind of the speaker" (6.53). If we could explore this idea, at the core is the view that by listen and understand the other feelings, we could improve our social relations, hence, a path to socially elevate our moral feelings to the highest virtue possible in the community.

4] This feeling of dignity in my view is what we lack in the state of nature. It could be explored that at the core of the concept of "unsocial sociability" as elucidated by Kant is the idea that what underlies our actions based on self-love is the logical absence of the concept of dignity (humanity). Consequently, we act based on our most basic instincts, hence on our own ends. However, once we in community, are recognized in terms of dignity and respect, which essentially signifies recognizing each other's humanity, we can curb our self-love interests. By sharing feelings, thoughts, and cognitive capacities, we can act empathically. This empathic stems from the fact that all of us are experiencing ourselves as free agents, but in Nietzschean terms.

But how can the ideas of dignity and respect empower people? I suggest that this empowerment must occur through a force that "becomes creative and ordains values." In other words, as Owen has pointed out, people need to experience themselves as agents. This can be achieved through the empowerment of dignity and respect that we project onto them. Therefore, this re-evaluation of the duties of virtue towards other human beings through the empowerment of their "ressentiment forces" by the feelings of dignity and respect can help them maintain their dignity. In this way, they will experience themselves as free agents and can overcome the asymmetrical relations of domination. Hence, there is no need to invoke a principle of empathy as the sole motivating force; it is only the foundational force that moves us through an empowerment process. Returning to the earlier example of people crossing the Darien jungle, we could be of more assistance by showing them dignity and respect[5] than by merely experiencing a sense of empathy.

In summary, while empathy can serve as the foundation for understanding the needs of others, it is insufficient on its own. The virtues of duty towards others, particularly dignity and respect, have the potential to empower individuals to reclaim their dignity, reassess themselves as free agents, and overcome social and economic asymmetrical relations. This internal reorganization of strength enables them to act as empowered agents, fostering a sense of independence within the community. However, this is contingent on a political commitment and a clear moral structure at both the individual and collective levels.

Allan.gonzalez.estrada@una.ac.cr

REFERENCES

Aurelius, Marcus. 2008. *Meditations*. General Editor, Kuund Haakonssen. Liberty Fund, Indianapolis.

David Owen. 2009. Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency. *Nietzsche on Freedom and Autonomy,* edited by Ked Gemes and Simon May, 197-222. Oxford: Oxford University Press.

——. 2017. Freedom as Independence. In *Nietzsche's Engagements with Kant and the Kantian Legacy*, edited by Marco Brusotti, Herman Siemens, João Constâncio, Tom Bailey, Maria João Mayer Branco, and Katia Hay, 137-58. London: Bloomsbury.

Joe Saunders. 2023. The Self, Love and the Other: Thoughts on Nietzsche, Kant and Owe. *Public Reason* 14(2)-15(1): 19-28.

Kant. Immanuel. 2018[1792]. *Religion within the Boundaries of Mere Reason*. Translated and edited by Allen Wood and George di Giovanni. Cambridge University Press. (RGV)

5] As long as 'dignity and respect' are expressed in political terms, it is essential to imbue classical notions that seek to establish morality with tangible content. In simpler terms, when a political entity declares that it recognizes the dignity and respects the lives of forced migrants (or other vulnerable populations), it implies a commitment to transfer power in the form of asylum, housing, health, and so forth.

———. 2017[1797].*The Metaphysics of Morals*. Edited by Lara Denis. Translated by Mary Gregor. Cambridge University Press.

Marias, J (ed. and trans). 1951. *Politica*. Madrid: Instituto Estudios Politicos. Edicion Bilingüe.

Smith, Adam. 2002. *The Theory of Moral Sentiments*. Edited by Knud Haakonssen. Cambridge University Press.

Public Reason 14(2)-15(1): 35-56

Moral Development, Repentance, and Self-Affirmation[1]

Paula Satne
University of Leeds

Abstract. This article engages closely with David Owen's 'Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency.' Owen argues that Kant tried, but ultimately failed, to resolve the tension between law and love that is characteristic of European modern philosophy. This is because Kant takes a 'highly critical stance to self-love throughout his moral philosophy' since he conflates self-love with psychological egoism and sees it as 'opposed to morality as a threat, a challenge, a danger…' Owen articulates Nietzsche's main objections to the Kantian opposition between self-love and the moral law and argues that Nietzsche has a rather greater claim than Kant to have resolved this tension. In this article, I explore whether some of the arguments developed in Owen's text represent serious challenges to Kant's position and develop a reading of Kant's ethics which can answer most of the challenges. First, I argue that Kant is committed to a version of the 'agency free will' model. I show that, on the Kantian model, one constitutes oneself as self by becoming what one is through reflecting about one's ability to live up to what one considers one's fundamental commitments. Second, I argue that on this model, there is no universal list of maxims or commitments that can be *fully* specified in advance for all agents in all times. Each person's project of moral transformation is personal and ultimately shaped by features of their own psychology and personal history as well as features of their social circumstances. Third, I argue that the Kantian project of moral transformation, which involves repenting immoral maxims, can also be understood as a form of flourishing and self-affirmation, which includes the pursuit of happiness as one of its components. On this reading, far from endorsing an 'ascetic ideal' of the moral agent, Kant embraces an ideal of the human life in which there is significant space, and even a duty, to pursue pleasurable endeavours. However, although morality and happiness are not intrinsically incompatible, they are likely to conflict under unstable and unjust external conditions. Thus, the extent to which we can live a completely morally good life is not independent of social and political conditions. Ultimately, in the Kantian picture, the task of moral self-improvement also has a social dimension: we must avoid being complicit with the injustices of our time and instead work towards overturning them. It is in fair (or at least fairer) social and political conditions that we can pursue both moral and personal ends and flourish.

Keywords: moral development, repentance, self-affirmation, moral transformation, Kantian forgiveness, happiness, flourishing, guilt, radical evil, Kant, Nietzsche, ethical and moral agency.

I. INTRODUCTION

This article engages closely with David Owen's *Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency*. Owen argues that Kant tried to resolve the tension between law and love that is characteristic of European modern philosophy by distinguishing morality into a domain of right and a domain of virtue. However,

1] Special thanks to Sorin Baiasu and Jakub Szczepanski for inviting me to contribute to the Rousseau Conference of the Standing Group on Kantian Political Thought 2022. To Jakub Szczepanski for commenting on this work, David Owen for his wonderful Rousseau lecture and Sorin Baiasu, David Owen, Matt Bennett, Allan Gonzalez, Sacha Mudd, Joe Saunders, Jakub Szczepanski and Franceso Testini for their thoughtful questions and comments.

Kant did not manage to fully resolve this tension. Although Kant managed to stress the importance of both self-respect and respect for others and make room for the love of others through the wide duty of virtue to promote their happiness, ultimately, he fails to provide a plausible account of true self-love. This is because Kant takes a "highly critical stance to self-love throughout his moral philosophy" since he conflates self-love with psychological egoism and sees it as "opposed to *morality* as a threat, a challenge, a danger…" (2009, 197). Owen articulates Nietzsche's main objections to the Kantian opposition between self-love and the moral law and demonstrates that Nietzsche "has a rather greater claim than Kant to have resolved the tension between love and law that is characteristic of European modern philosophy" (220). In this article, I identify two main challenges that Nietzsche's view on ethical agency (at least as it is interpreted by Owen) poses to Kant's theory of moral agency (at least as it interpreted by me). My aim is to comment on Owen's interpretation of Nietzsche as I understand it. I don't have enough knowledge of Nietzsche to evaluate this interpretation or to challenge it. Nor is my aim to argue that Nietzsche's views of ethical agency are better than Kant's model of rational agency or vice versa. Instead, my aims are to explore whether some of the ideas developed in Owen's text represent serious challenges to Kant's position and to develop a reading of Kant's ethics which can answer most of the challenges. I will proceed as follows: In section II, I start by offering a brief reconstruction of Kant's account of rational agency and moral development. In section III, I examine Nietzsche's model of agency free will, which is expressed in his notion of the 'sovereign individual' and argue that it has significant similarities with Kant's account. In sections IV, V, and VI I address the two main challenges that Nietzsche's position poses to Kant according to Owen and argue that Kant's moral philosophy has the resources to answer them. Section VII briefly reflects on the place of flourishing in Kant's philosophy by way of conclusion.

II. KANT'S ACCOUNT OF RATIONAL AGENCY AND MORAL DEVELOPMENT

In previous work, I have developed an account of Kantian forgiveness by appealing to Kant's theory of rational agency, his theory of imperfect duties, and his theories of radical evil and moral development. Here I offer a reconstruction of Kant's theories of moral development and rational agency with the aim to then be able to compare them to Nietzsche's position as reconstructed by Owen. For Kant the will is practical reason, that is, a faculty of acting through the conception of a principle. Kant distinguishes two types of principles. Objective principles hold for all rational beings and instruct us how we ought to act, and for finite beings like ourselves, take the form of imperatives (categorical and hypothetical) (G 4: 413).[2] Subjective principles are maxims, that is, self-given principles of action that hold

2] Following standard practice, references in the text to the works of Immanuel Kant use the numbering from the Akademie edition, with the following abbreviations of titles: *Groundwork of the Metaphysics of Morals*: G; *Metaphysics of Morals*: MM *Lectures on Ethics*: LE; *Critique of Practical Reason*: CrP; Religion within the Limits of Pure Reason: Rel.

only for the subject (G 4: 422). A central feature of Kant's theory of rational agency is that agents act on maxims. Maxims typically express the reasons that motivate an agent to act and thus they express the agent's rationale and justification to perform that action. More formally, a maxim is a principle that connects some generic description of circumstances (taken broadly to include the inclinations and purposes of the agent) with some generic description of an action type that the agent takes these circumstances to warrant. Kant claims that we are sometimes uncertain of our own motivations (G: 4407; Rel G 6: 20), which means that we are not always explicitly or consciously aware of the maxims that we adopt. Maxims can be adopted tacitly, implicitly and, in many cases, retroactively. However, as maxims are a product of our freedom and principles, for which we are responsible, we can and should become aware of them through reflection (Korsgaard 1996). Moreover, Kant is clear that the *first command of all duties to oneself* is '*know* (scrutinize, fathom) *yourself* …in terms of your moral perfection in relation to your duty' (MM 6: 441). Thus, a sign of rationality, and an expression of a good character, is a willingness to take responsibility for one's actions by examining and revising one's maxims. Thus, reflection about one's maxims and actions (self-reflection) are a central feature of Kant's theory of agency.

Another central aspect of Kant's account is that since the default position of the will is a position of 'radical evil' (Rel 6: 32), it is ethically necessary, and therefore it must be possible, to overcome this evil (Rel 6: 66-67). To overturn evil is to take on the task of becoming virtuous in the sense of acquiring a good *Gesinnung* (meta-maxim): to make one's commitment to the moral law unconditional, which requires a 'revolution of the heart' (Rel 6: 47, 51) or inversion of the incentives, consisting in subordinating the maxim of self-love to the moral law. The revolution of the heart provides the rational framework that allows a person to abandon her immoral maxims. This is because lower-order maxims are rationally justified by higher-order maxims, so a fundamentally good person has no grounds of justification for more particular immoral maxims. For Kant, moral development involves reforming our characters by becoming aware of our maxims and getting rid of those maxims that on reflection we do not fully endorse. Thus, moral development requires self-reform. Moreover, I interpret the reorientation of one's will as requiring a single revolutionary act. However, after (or during[3]) the revolution there is still more progress to be made (Rel 6: 47-48; 66-67) because although the revolution rules out 'vice,' that is, the principle of deliberately violating a duty (MM 6: 380), it does not rule out

3] Kant says that the choice of evil *Gesinnung* (i.e., the fundamental disposition) is an 'intelligible deed' that does not occur in time (Rel 6: 31) which is sometimes taken to imply that the choice of a good *Gesinnung* is equally timeless. The challenge is then to explain the relation between the timeless revolution and the gradual process of moral self-improvement involved in a project of moral self-transformation. Some authors solve the problem by arguing that the process of overcoming radical evil occurs simultaneously with the process of improving the morality of one's maxims, a process that takes place over the course of a person's life (Sussman 2005, 173; Korsgaard 1996, 181) while others are willing to allow that there is a sense in which we can say that the revolution can be placed at a moment of a person's life (Drogalis 2013, 166). The point I am making here about moral progress and the ongoing struggle to achieve moral perfection aims to be compatible with both readings (see Satne 2016, 1043).

'impurity' and 'frailty' (MM 6: 408). The striving towards virtue, which I identify with the possession of a good meta-maxim,[4] requires constant (endless) progress and a continued effort to approximate an (unattainable) ideal of holiness (MM 6: 409 and 390), understood as the aim of acquiring a fully reliable and pure form of moral motivation.

This process of self-reform involves the abandonment of our immoral maxims and thus it necessarily involves repentance for our immoral acts.[5] I understand the Kantian notion of repentance primarily in intellectual terms as simply the commitment to abandon immoral maxims and become a better person as part of an ethical project of self-reform. Although the intellectual aspect of repentance is the primary aspect of the notion, Kant also states that repentance has a painful, affective aspect: "repentance is a painful sensation which is brought about by a moral attitude" (CrP 5: 98). This is the painful feeling of guilt which usually accompanies the judgment that we are morally guilty in the sense of recognising that we acted wrongly (MM 6: 399). Kant, however, rejects the notion of repentance as a form of self-punishment or chastisement (LE 27: 464) and instead maintains that "inner contrition for our offences and the firm resolve to live a better life" are the only things that are truly helpful to ourselves and others (LE 27: 464). I call this account 'minimalist' because although repentance might include pain, it is not the pain that is the motivational aspect of the state.[6]

This characterisation of Kant's account of rational agency and moral development has been brief[7] but I will continue to develop these and other aspects of Kant's position throughout this article as I discuss the challenges presented by the Nietzschean position as interpreted by David Owen.

III. NIETZSCHE'S SOVEREIGN INDIVIDUAL

Owen presents a picture of Nietzsche's account of ethical autonomy and self-respect by providing an analysis of the figure of the 'sovereign individual' as presented in the second essay of *On the Genealogy of Morals*. Nietzsche presents:

> an image of [ethical] *maturity* articulated in terms of *individuality*, where such individuality is linked to *autonomy*, that is, being able to impose binding norms on oneself ('the man with his own independent, enduring will, whose *prerogative it is to promise*'). This figure represents, it seems, the concept of the autonomous individual who is not bound by moral rules as customary constraints, but as *the freely endorsed commitments through which he gives expression to his own character* (204).

4] I defended the claim that the possession of a good meta-maxim (*Gesinnung*) amounts to the possession of a virtuous character in Satne 2013 and 2016.

5] Kant explicitly says that if there is no excuse or reparation for an offence, the wrongdoer should show "contrition" and "regret." But if "the injured party is not content with that, then it does a man honour if he offers an apology", noting that "it is not degrading to apologize" (LE 27: 435).

6] See Satne 2021 for a discussion of the relation between the notions of guilt and repentance in Kant.

7] This interpretation is developed in more detail in Satne, 2016.

Nietzsche equates the sovereign individual with the "liberated man who has the *prerogative* to promise, [the] master of *free* will…" (GM II, 2). Nietzsche's equation of the sovereign individual with 'free will' can represent an obstacle to Owen's argument that the sovereign individual represents Nietzsche's account of ethical maturity since Nietzsche rejects the doctrine of free will 'in the superlative metaphysical sense,' with its underlying commitment to moral accountability, responsibility, and retribution and its appeal to the idea of the agent as *causa sui*. To overcome this obstacle, Owen appeals to Ken Gomes's distinction (2009) between the '*deserts* free will' model and the '*agency* free will' model. Owen maintains that Nietzsche rejects the model of *deserts* free will**,** which is directed to the justification of blame and is represented in the form of debates over free will and determinism, but he accepts an *agency* free will model, which is directed to the question of what distinguishes mere doings (events) from deeds (actions) and relates to debates over the conditions which must be met for one to be an agent, for one's doings to be deeds. Arguably, Kant is committed to a complex version of the deserts free will model. However, I won't discuss this aspect of Kant's position here. Instead, I will be arguing that Kant is also committed to a version of the agency free will model.

The sovereign individual possesses self-mastery in the sense that he has full competence to represent himself to the rest of the world as someone who can maintain a commitment or a promise. At the opposite pole, is the liar, a mere *wanton*, who is incapable of keeping his word. However, since arguably the majority of socialized individuals are not wantons, Owen attempts to explain what is *distinctive* about Nietzsche's account of agency. One distinctive aspect of the sovereign individual is that in this model, one's deeds are seen as criterial for one's intentions. This means that the sovereign individual is not only committed to fulfilling his promises, but he does indeed fulfil them. The only impediments to the realisation of his commitments, which are compatible with self-mastery, are physical or normative impossibilities. Since it is impossible to have mastery over faith in general, a degree of prudence in one's commitment-making activity is also recommended. Owen stresses the importance of upholding one's word "even against fate," which involves taking a stance which "acknowledges and affirms the fatality of one's agency" in the sense of accepting "the damage incurred when one's commitment cannot or must not be kept" (207). This aspect of Nietzsche's position can be aligned with Bernard Williams's analysis of the claim 'I did it.' Contrary to the 'morality system', which does not pay attention to the significance of the 'I did it', "the sovereign individual is one for whom the thought *I did it* has ethical purchase and salience" (208). The ethical significance of the 'I did it' is highlighted in Williams' example of Ajax's suicide. After a deranged act, Ajax commits suicide out of shame, acknowledging the significance of what he did: "being what he is, he could not live as a man who had done these things…" (Williams 1993, 73). I will analyse the significance of 'I did it' for Kant's position in section VI.

The second distinctive aspect of Nietzsche's position is that the sovereign individual is able to master the norms constitutive of social practices. While the success conditions of promising can sometimes be specified externally, promises often also have internal conditions that cannot be fully specified in advance. In this case, the nature of my intention is revealed in the way that I keep the promise. For example, the nature of my intention when I promise to 'love and honour you until death do us part' is only revealed in the way I treat you throughout our lives together and requires an understanding of the social practices through which we act (i.e., in this case the social practice of marriage). Promising is thus an essential part of *becoming what one is* and of *constituting oneself as self*:

> It is only through commitments characterized by internal success conditions that one discovers what one's intentions are or…acquires self-knowledge concerning what one is, what is necessary to oneself and what is not…the process of *becoming what one is* requires self-knowledge…[about] one's commitments and the issues that arise in the practical lived working out of what one has committed oneself to in making these commitments… (210).

We can see how some of the features of the model of agency here sketched are also implicit in Kant's conception of rational agency as sketched in section II. Recall that Kant claims that we act on maxims and that these maxims are always opaque to us, but that we nevertheless have a duty to reflect about our true motives. Arguably, given our tendency to deceive ourselves about our true motives and the inherent opacity of our maxims, any possible knowledge about one's commitments can only be acquired through reflecting about the success conditions internal to our intentions (maxims in Kantian parlance). This, of course, will involve knowledge of the practical issues that arise when one tries to live up to one's maxims. Thus, on the Kantian model of agency, one also constitutes oneself as self by becoming what one is through reflecting about one's ability to live up to what one considers one's fundamental commitments. The commitment to a project of moral self-improvement is also a form of constituting oneself as self and it also requires self-reflection, although complete self-knowledge is, for Kant, beyond human reach.

There are some important similarities between Nietzsche's account of ethical autonomy and Kant's model of moral agency as the following passage clearly shows. For Nietzsche, freedom demands:

> [t]he ability to take one's virtues and oneself as objects of reflection, assessment and possible transformation, so that one can determine who one is [...] to take ourselves as potentially free requires that we are not merely bearers of good qualities but *self-determining beings capable of distanced reflection* (Guay 2002, 315, cited at 211, *my emphasis*).

The Kantian moral agent is conceived precisely as an agent who is capable of taking distanced reflection and self-determination. Owen acknowledges that Nietzsche's position shares with Kant a view of autonomy as the ground of self-respect. He draws

from Nietzsche's discussion of freedom in *Twilight of the Idols* and argues that there Nietzsche also identifies freedom with standing surety for oneself but he adds the claim that, given our exposure to luck, self-mastery is a condition that can never be fully and finally achieved:

> the condition of self-mastery must be construed as an *ongoing achievement* or, to put it in another way, a *continual process of struggle engaged with both the material of oneself and the circumstances of one's agency*. [...] [T]he sovereign individual, the ethically autonomous person [i]s one who is engaged in a process of struggle which continues through the course of one's life (212, *my emphasis*).

As we have seen, Kant also emphasizes the ability to reflect on one's maxims and assess one's actions as a central aspect of autonomous agency and he takes the possibility of radical transformation of the self as a central aspect of our freedom. Moreover, for Kant this process of transformation involves a struggle which continues through the course of one's life. The struggle of the Kantian moral agent can arguably also be seen as a struggle with the material of oneself understood as both one's maxims (commitments) and our ability to shape our inclinations and emotions in line with those commitments. Moreover, although this is not often recognised, in section IV below I will argue that the Kantian moral agent must also struggle with the circumstances of their agency and the injustices of their time. Having shown the similarities between Kant's and Nietzsche's accounts of free agency, in the next two sections I consider the two central challenges that Nietzsche's position poses to Kant's account, the second challenge being Owen's central preoccupation.

IV. NIETZSCHE'S FIRST CHALLENGE

So far, I have been able to show that both Kant and Nietzsche endorse versions of the agency free will model and although there may be differences between the two, there are also some significant similarities. However, Owen's reconstruction of the Nietzschean position poses two important challenges to the Kantian model.[8] The first challenge arises because, according to Owen, Nietzsche provides an *ethical* model of agency, while Kant endorses a *moral* model of free agency. According to Owen, Nietzsche is articulating a view of *ethical* autonomy that contrasts sharply with the ideal of *moral* autonomy expressed in Kant and which Nietzsche takes to be representative of 'morality' (2009, 210). The problem for Kant is the central role that the categorical imperative plays in his moral philosophy since this implies that if:

> I find that the maxim of my action cannot be universalized without contradiction, I have identified an absolute prohibition, an unconditional 'I will not'. I have, in other words, stopped short at a formulable instruction that might be fully obeyed by anyone… The spirit… has gone missing without trace (Ridley 2009, quoted in 210).

8] Owen's texts also raise other (minor) criticisms to Kant's position, which I address as I proceed.

On this reading, the Kantian moral agent is obliged to comply with a list of 'I will not's' that can be specified in advance and independently of the way in which commitment to them is executed. The characteristic error of 'morality' which is enshrined in Kant's moral philosophy is a failure to acknowledge the expressive character of human agency and thus Kant is unable to account for the nature of "human freedom as an unformulable process of self-legislation" (210).

Kant argues that moral requirements impose constraints on our maxims of action and thus give rise to moral obligations. Thus, Kant certainly can be read as imposing a list of 'I will not's' on agents. However, I argue that this is somehow a simplistic reading of Kant's ethics. In fact, I do not believe that a list of maxims or commitments for each Kantian agent can be fully specified in advance. It is true that moral demands are universal. But morality is not simply an issue of endorsing a list of 'I will not's'. Often (usually) we need to think much more carefully about what is morally required from us. Each person's project of moral development has both individual and collective (historically situated) aspects. Moral development has an individual dimension because the inclinations that represent temptations to immorality vary for each person since they depend on features of our psychology and individual histories. Moral development involves reforming one's character by revising one's maxims and this is going to involve something different for each one of us. The key task for each agent is to figure out how can their maxims, which are subjective principles of action and thus relate to all aspects of an agent's life (including their work, personal projects, aims and aspirations, personal relationships, and community life, among many other areas aspects), embody universal values, i.e., respect for the dignity of persons.

Moreover, for Kant, wrongdoing is in broad terms a tendency to either act on subjectively valid motives while recognising that they do not provide justification for one's actions (weakness) or, more seriously, a tendency to take subjective valid motives as having more objective force than they really have in the sense of being justified from the standpoint of others (depravity). The theory of the *Gesinnung* goes some way to explain how this happens, since the agent adopts self-love as her fundamental maxim, subjective valid motives are seen as having objective justification. However, given that the moral law commands unconditionally and that we are always aware of moral obligation since consciousness of the moral law is a basic 'fact of reason' (CrP 5:31), what is more difficult to understand is why we have this predisposition to adopt self-love as our fundamental maxim.[9] One way of understanding Kant's position is to see self-love as sustained by ideology or rationalizations. Andrews Reath expresses the point as follows:

9] In fact, Kant's theory of radical evil is notoriously complex and is considered by many as one of the most controversial and difficult aspects of Kant's moral psychology. Here I am glossing over some of these complexities in the interest of clarity of exposition and because the points I want to make here do require that we engage with these complexities. For an overview of some of the main difficulties associated with Kant's radical evil thesis, see Morgan 2005, 63-65.

> The different tendencies that this description [of taking merely subjectively valid motives to have objective justification] fits might share the feature of being sustained by some set of false or impoverished beliefs. These could range from beliefs about one's motives or the relevant features of one's situation, to beliefs about practical reason or the moral capacities of the self [...]. [Thus,] the influence of self-love on the will is sustained by an ideology of sorts, which enables individuals to view their maxims as objectively acceptable reasons (2006, 21).

The ideologies that sustain self-love are often developed collectively (although sometimes can also be formed by a sole individual) and thus they can vary both historically and geographically. This means that temptations to immorality (this time in the form of false beliefs) and the challenges that each person must overcome in their struggle against radical evil often vary from person to person. Myisha Cherry (2021) provides the example of Christian Picciolini. Christian Picciolini was a member of a violent neo-Nazi organization, which he eventually rejected. He became the author of *Breaking Hate* and created a non-profit organization to help other disengage from hate groups.[10] On a Kantian reading of Picciolini's transformation, we can see him as engaging in a personal project of moral development, which includes his own personal set of maxims, dealing with the particular obstacles that he had to overcome given his personality and social circumstances. Ultimately, there is no universal list of maxims or commitments that can be fully specified in advance for all agents in all times. Each person faces their own challenges and has to attempt a transformation of different maxims.

It is also important to note that some of our obligations are context dependent in the sense that each society has its own structural injustices, which give rise to different moral challenges and, in some cases, collective obligations. Each époque must interpret what morality demands of them at that time and place. Here I have in mind cases in which we have no way to act rightly because we live in a structurally unjust state such that, despite our best efforts, we cannot but be implicated in wrongdoing. This type of case has recently been discussed by Lucy Allais (2015) and Garrath Williams (2018). Williams focuses on our complicity with climate change especially in developed societies. Allais focuses on the moral dilemma of whether or not to give money to a beggar. Starting from the thought that just public institutions must uphold subsistence rights in order to prevent relations of dependence and domination, Allais shows that societies that fail to provide poverty relief are structurally unjust and often leave their citizens with no way of acting rightly. However, on closer analysis, even though we appear to be left without options, we are often complicit with that structural injustice because we cooperate with and benefit from it and thus we bear *some*[11] responsibility for it. As Allais eloquently puts it:

10] Cherry 2021, 172, fn. 12.

11] There are degrees of power, responsibility, and complicity and of course some are much more implicated than others.

> On Kant's analysis, our daily confrontations with beggars confront us with *structural injustice in which we are implicated*. In this situation, there may be nothing you can do to relate to the beggar perfectly rightfully…[T]his analysis …suggests that, although *you do not do wrong if you do the best act available to you, the extent to which you can live a completely morally good life is not independent of the conditions of the society you live in*. Living in an unjust state means you can find yourself in situations in which there is no morally unproblematic alternative [...] If there are genuine beggars, then *there is structural injustice of a sort which makes it impossible for you to relate rightfully to these individuals in individual encounters*. (2015, 770-71, my emphasis).

Allais says that living a completely morally good life is not independent of social conditions. This is somehow surprising because it is usually thought that Kant's ethics is an ethics of autonomy and that the extent to which you can live a completely morally good life depends on your maxims or principles of actions and ultimately your individual character. However, Williams' and Allais' analyses show that there is a social dimension to our attempts to live a moral life and thus the continuous struggle to live a moral life also requires engagement with the circumstances of one's agency. Thus, we are (partially) responsible for the injustices of our time and attempting to address these injustices often requires acknowledging our complicity and embarking on collective action. Here again the task ahead of us is one of moral and social transformation by committing to an individual and collective project of reflection and reform. As noted by Allen Wood, moral progress and the struggle against radical evil "can be effective only if it is carried out through an ethical community" (Wood 1999, 332).

Thus, the Kantian moral agent is not simply committed to a list of 'I will not's'. On the contrary, each person's project of moral development is personal and shaped by features of their own psychology and personal history as well as features of their social circumstances. However, it is true that the Kantian process of self-legislation is unlikely to be completely unformulable. To the extent that our maxims must express respect for the dignity of persons, and given that dignity expresses an absolute worth, I see no advantage in a process of self-legislation that is completely unformulable.

V. NIETZSCHE'S SECOND CHALLENGE

The starting point of Owen's article is Aristotle's distinction between 'true' self-love and 'pseudo' self-love in Book IX of the *Nicomachean Ethics*. True self-love is "an acquired relation to self that consists in appropriating to oneself (committing oneself to the service of) what is noblest by submitting to the rule of reason and, hence *identifying one's well-being with living a virtuous life*" (199, *my emphasis*). In contrast, pseudo self-love simply involves the self-interested pursuit of pleasure and the identification of happiness with the satisfaction of one's irrational appetites. Owen claims that despite the plausibility of such distinction Kant denies it any place in his moral philosophy, attributing this denial to Kant's view that inclinations are always both the product

of luck and are exposed to luck and that, therefore, they cannot be conceived as the objective basis for the ascription of moral worth. For Kant, respect for the moral law is the "basis of the disposition of self-respect (as well as that of respect of others) that characterizes the morally autonomous agent" (202). By denying the Aristotelian claim that the moral agent is characterised by a disposition of true self-love, Kant fails to fully resolve the tension between love and law that informs the modern philosophical tradition. Owen's conclusion is that Nietzsche's ethics has a "rather greater claim than Kant to have resolved [this] tension" (220).

According to Owen, for Nietzsche: "true self-love is the [acquired] disposition to value what is noble" (213). Nietzsche's conception of nobility in *The Gay Science* has two main elements: (i) *amor fati*, which involves "learning to love the piece of fate that one is by coming to see what is necessary in and to oneself as a new 'indescribable beauty'" (213) and (ii) the doctrine of eternal recurrence as "the lover of one's life 'who no longer want anything better from the world than it and it again' (GS 334)" (213). For Owen, it is the disposition to *amor fati* that is the silent disposition of *true* self-love' (213). Nietzschean self-love is a form of artistic interpretation through which one comes to see oneself as a 'new and indescribable beauty' (GS 334). Nietzsche exhorts us to embark on the ethical work needed to '*become what we are*—human beings who are new, unique, incomparable, who give themselves laws, who create themselves!' (GS 335) and "to 'give style' to one's character- a great and rare art!" (GS 290).

In Owen's interpretation, to submit to one's taste or, in the case of ethics, to submit to one's own law involves a process of self-creation and self-formation, which involves being bound by constraints related to "the intentions of the artist which, in the case of self-artistry, are comprised of one's evaluative commitments" (215). Submitting to the demands of one's own law does not mean that the relevant constraints are not objective in the sense of being independent of the relevant subject. To commit to the value of scientific activity, for example, may be an expression of one's taste but "the value of scientific activity entails that the constraints to which one is subject are those norms and necessities that govern the practice of scientific activity. Hence it follows that one can only give style to one's character by submitting to the norms and necessities that govern the relationships, practices, ideals, etc. to which one is committed" (215, fn. 28).

Owen reads Nietzsche as offering an aesthetic re-description of the sovereign individual and claims that:

> A preliminary view concerning the relationship between self-love and self-respect would thus be that valuing the disposition of *amor fati* broadly equates to valuing the will to self-responsibility (i.e., ethical autonomy) and, hence, that *self-love consists in valuing self-respect* (216, *my emphasis*).

In the case of self-artistry, the medium is the human material comprising oneself. Giving style to one's character involves surveying all the strengths and weaknesses that our nature has to offer (GS 290), removing non-necessary weaknesses and ugliness and

concealing and reinterpreting necessary weaknesses "so that one learns how to see it as beautiful, as a strength" (Ridley 2005, xiii, quoted in 217). Owen links this discussion of self-artistry to the discussion of intellectual conscience in GS 335:

> A necessary condition of the self-mastery exemplified by the sovereign individual is the exercise of intellectual conscience and, hence, truthfulness [...] with respect to what one is, what one is committed to and the requirements of these commitments. At its limit, this is an unflinchingly honest and realistic appraisal of oneself and the circumstances of one's agency… (217).

To maintain such a stance is difficult and requires strength. In fact, we need art as a counterforce to guard off nausea and suicide and to maintain our overall commitment to truthfulness (GS 107). Artistry, and the "local falsifications, projections, rationalizations, and illusions" that come with it accompany our search for truthfulness and are also part of valuing the disposition to *amor fati*, which constitutes true self-love (218). For Owen, Nietzsche's ethics seeks to dissolve the tension between self-love and self-respect by presenting *self-love as the disposition to value self-respect*.

Nietzsche's second challenge to Kant then is simply that:

> Kant's framing of this distinction [the modern distinction between morality and egoism] rules out self-love as the general orientation to *eudaemonia* (whether conceived in Aristotelian or Nietzschean terms) and, hence, fails to acknowledge the centrality of self-love to the achievement and maintenance of the disposition of *amor fati* (Nietzsche's version of eudaemonia) (219).

The objection is not that Kant fails to give a place in his moral system to self-love or happiness understood as the overall sum of one's inclinations. Owen claims that Nietzsche joins both "Aristotle and Kant in rejecting the life of pleasurable inclinations, of pseudo self-love" (219). The way I read the objection is that ultimately Kant does not have an account of flourishing.

Owen notes that Kant's rejection of self-love is bound up with his effort to immunize morality from luck (219). However, given Nietzsche's naturalism in ethics, he cannot possibly endorse this project. Thus:

> Nietzsche's strategy is exactly the opposite of Kant's; it is to build an acknowledgment and affirmation of our exposure to luck into our understanding of ethics and ethical autonomy. This strategy embraces self-love as necessary to the task of developing and maintaining an orientation to, and engagement with, the world that consists in *leading* our lives and not simply existing through the course of their duration (219).

I find the claim that self-love consists in valuing self-respect somehow difficult to (fully) understand (i.e., in what sense is this a form of 'love'?). But, in any case, the key point seems to be that on Nietzsche's conception of autonomous and self-respecting agency, the agent achieves a form of flourishing or eudaemonia through loving their fate (*amor fati*), a form of becoming and affirming oneself which involves self-creation, but which also requires commitment to truthfulness and self-artistry. The objection, as

I understand it, is that Kant's ideal of moral agency fails to constitute a form of genuine flourishing or *eudaemonia***.** The question that arises specifically for my interpretation of Kant is this: to what extent can it be said that the ongoing project of self-reflection and self-reform, which represents a fundamental commitment of the human agent, but which is also, to a certain extent, painful (because it necessarily involves the painful aspect of repentance) and a struggle, leads to a life of flourishing?

Considering the interpretation of Kant sketched in the previous section, we can say that Kant is committed to a version of the claim that submitting to one's own law involves a process of self-creation and self-formation which involves 'giving style' to one's character. Although, it would be stretching too far to say that we can find a reference to self-artistry in the Kantian process of self-creation (at least in Kant's ethical writings), there is a sense in which it can also be said that Kant is committed to a version of the claim that human beings are new, unique, incomparable, give laws to themselves and create themselves. Moreover, Kant also thinks that an unflinchingly honest and realistic appraisal of oneself and the circumstances of one's agency is morally required, although he thinks that we can never be sure that we have achieved total self-awareness. So, there is a sense in which we can say that the Kantian ideal also involves self-creation and self-reflection. Although Nietzsche's account seems to put more emphasis than Kant's on reinterpreting and ultimately accepting our weaknesses as beautiful, Kant, Aristotle and Nietzsche all stress the importance of self-reformation and critical appraisal of oneself (Owen, 2009, 200). So, the crucial questions really are: can the Kantian project of moral transformation also be seen as a form of self-affirmation? Can it be seen as a form of flourishing?

Given that the moral law is a *fact of reason* and that we experience it with reverence (respect), it is (relatively) easy to explain why we can be motivated to act from duty when moral demands conflict with inclination. However, the interpretation here defended implies a stronger claim: to be a moral agent is to be an agent who is fundamentally committed to the moral law as this involves commitment to an ongoing, long-term, ethical project of self-reflection and self-reform. In short, why should we adopt the moral law as our fundamental maxim? In a recent article, Anastasia Berg eloquently explains it thus:

> because moral law commands universally, properly recognizing its authority means recognizing its authority universally, or, over all of one's actions. In particular it means that an agent cannot be thought of as only occasionally acting from genuine recognition of the moral law. To allow oneself to occasionally deviate from the commands of the moral law is simply to refuse its authority as universally binding, which is to refuse its authority *tout court*. An agent is thus either committed to the supremacy of moral considerations in all matters, or, in effect, in none (2022, 178).

This explains why we must commit to the moral law unconditionally, and why there is no space for moral luck in Kant's system. However, although adopting the moral

law as our fundamental maxim is morally necessary, this still does not show how or whether such moral life amounts to a life of flourishing.

However, at this point, it is important to make some important clarifications about Kant's position and in particular the role that happiness plays in Kant's ethics and his theory of moral development. In order to make this argument, I will now complement my reconstruction of Kant's position by drawing from Alice Pinheiro Walla's recent book *Happiness in Kant's Practical Philosophy: Morality, Indirect Duties, and Welfare Rights.* Pinheiro Walla shows that the pursuit of happiness plays a much greater role in the personal and collective life of the Kantian moral agent than is commonly thought. Kant rejects the principle of happiness as the basis of human morality because he thinks that an adequate moral theory must be one that takes seriously the common understanding of morality as unconditional binding. For Kant, this means that "autonomy of the will is the sole principle of all moral laws" (CrP 5: 33). Kant rejects ethical theories which are based on the principle of happiness as heteronomous since they presuppose a prior object of the will to be authoritative and thus, they can only generate hypothetical, and hence, conditional imperatives. However, Pinheiro Walla shows that Kant's rejection of happiness as the foundation of morality should not be confused with a hostile attitude toward human happiness in general and an attempt to eradicate happiness from the moral life.[12]

Pinheiro Walla interprets Kant's claim that human beings have the end of happiness by natural necessity (G 4: 415-6) as following from Kant's conception of the finite *rational* will. Maxims make references to acts which are pursued for the sake of some end. To act in the world, we also need to adopt *ends,* which constitute the necessary material aspect of willing. Kant claims that adopting an end analytically implies the commitment to take the means for its realization (G 4: 417). The matter of the will is provided by our inclinations understood in broadly hedonist terms as desires to obtain pleasure. The inclinations provide the matter for the adoption of our non-moral ends. However, since having an end necessarily involves one's commitment to its realization, in order to realize our non-moral ends and satisfy some of our inclinations, we must attempt to form a determinate material and realistic conception of our happiness. As Pinheiro Walla puts it "because we have a plurality of ends which are incompatible with each other or must be realized in different times, agents are confronted with the task of forming a conception of the ends that constitute their happiness, in a more or less coherent hierarchy" (2022, 10). While each agent's material conception of happiness would differ and agents would also differ in their ability to form a coherent model of happiness and live up to it, the "necessity" of forming a determinate and realistic conception of one's own happiness understood as well-defined sum of compatible ends arises from the structure of finite willing itself and it is not imposed externally from our non-rational animal nature.

12] This argument is made in Chapter 2 of the book.

Kant also claims that happiness cannot be the *natural* end of finite beings with the capacity of practical reason (G 4: 395). Adopting the perspective of the teleological principle of natural efficiency, Kant argues that we cannot assume that happiness is the highest end of nature for humanity since instinct seems to be a more efficient capacity for the attainment of our happiness than reason (G 4: 395). However, the fact that our reason is not merely contemplative but also has a practical use provides ultimate evidence for the claim that happiness cannot be the highest end of humanity. The only end that reason is best to promote "by itself" is the good will, which must be seen as the natural end for human beings. Pinheiro Walla shows that it is precisely because happiness is not our natural or whole end that incompatibilities between happiness and morality are merely contingent and not intrinsic to morality. If we took happiness to be the fundamental principle of the will, then the possibility of morality would be excluded as a maxim of subordinating morality to happiness is evil (Rel 6:31) whereas subordinating happiness to morality does not exclude the possibility of happiness. Having morality as our natural end allows us to see the natural and social evils that plague humanity, not as something that we have no hope to improve, but as presenting us with a moral task, that is, as something we have a duty to address. Paradoxically, Pinheiro Walla notes, the recognition that morality is humanity's supreme good can "bring us closer to contentment than making happiness our 'whole end'" (2022, 22).

Kant claims that we have an indirect duty to promote our own happiness (G 4: 399). Kant's explicit rationale for the duty is to make us less susceptible to temptations to immorality arising from an unhappy life. The underlying idea seems to be that a life of continuous and unbearable discontentment would make the pursuit of morality more difficult and perhaps even impossible (Pinheiro Walla 2022, 80). Kant argues that we have a duty to promote our long-term happiness and to make sacrifices for the sake of our long-term well-being. Pinheiro Walla explains Kant's position by arguing convincingly that Kant makes an implicit distinction between subjective and objective happiness. Objective happiness refers to "basic ends of our animal nature which have an impact on our moral integrity and thus can be commanded in case of neglect," while subjective happiness simply refers to what we happen to desire (2022, 82). Pinheiro Walla argues that securing one's objective happiness can become the object of a direct duty "presumably when the agent feels no inclination to pursue her own happiness and the neglect of her wellbeing has moral relevance" because it either has an impact on her capacity for moral agency or her moral integrity (2022, 81). Thus, the indirect duty to promote one's happiness can qualify as a direct duty under specific circumstances.

Pinheiro Walls's insightful analysis has important implications for my interpretation of Kant's theory of moral development as requiring an agent's commitment to a project of self-reflection and self-reform. Given human cognitive limitations, happiness would always involve an element of indeterminacy (G 4:418). However, it is rationally necessary for finite beings like us to attempt to form a determinate (as much as we can) and realistic conception of our own happiness understood as well-defined sum of our

compatible ends. Moreover, we have an indirect (and sometimes even a direct) duty to promote our own happiness. As we have seen, actual perfection is an unachievable task for finite beings. Pinheiro Walla reminds us of Kant's warning against the moral fanaticism involved in striving to achieve perfection in our finite lives at all costs, often attempting to do so by searching for spurious opportunities to act morally and ultimately "turn[ing] the government of virtue into tyranny" (MM 6: 409).[13] Here we can conclude that the pursuit of happiness, which nevertheless cannot be reduced to the contentment that follows from complying with morality, constitutes an important aspect of each agent's personal project of moral self-improvement.

With this analysis in place, we can see that although Kant rejects the principle of happiness as the basis of morality, it is inaccurate to say that Kant adopts a 'highly critical stance towards happiness throughout his moral philosophy.' Although it is true that Kant saw the tendency of the human heart to make happiness our supreme principle as the source of the radicality of evil in the human condition, it is nevertheless inaccurate to say that Kant saw happiness *per se* 'as a threat, a challenge, a danger' to morality. On the contrary, it is precisely because morality is the natural end of human beings that morality and happiness are not necessarily incompatible and that there can be space for the pursuit of happiness in human life, albeit a pursuit which is subordinated to the primacy of morality. Perhaps more surprisingly, although prioritising our own long-term well-being requires short-term sacrifices, Kant's conception of happiness should be understood in broadly hedonist terms. This means that, far from endorsing an 'ascetic ideal' of the moral agent, Kant can embrace an ideal of the human life in which there is significant space, and even a duty, to pursue pleasurable endeavours. It turns out that Kant's ethics is much more welcoming to the pursuit of pleasure than either Aristotle's or Nietzsche's accounts.

The pursuit of happiness has value and plays a central role in the life of a Kantian moral agent but, as we have also seen, morality has priority over happiness and, when the two conflict, Kant's position entails that we must sacrifice happiness for the sake of morality. Thus, there is a seemingly tragic aspect of Kant's moral theory, as in some difficult cases, it may even demand the complete sacrifice of our happiness for the sake of complying with our (perfect) duties:

> Ask a [man] whether, if his prince demanded, on pain of the same immediate execution, that he give false testimony against an honourable man whom the prince would like to destroy under a plausible pretext, he would consider it possible to overcome his love of life, however great it may be. He would perhaps not venture to assert whether he would do it or not, but he must admit without hesitation that it would be possible for him. He judges, therefore, that he can do something because he was aware that he ought to do it and cognizes freedom within him (CrP 5: 30).

13] Pinheiro Walla 2022, 113.

Although the conflict between morality and happiness can certainly be tragic, we should not see this tragic conflict as an inherent aspect of the human condition. As we have just seen, since morality is the end of human beings, this means that morality and happiness are not intrinsically incompatible. Here again Pinheiro Walla makes an important observation: often whether morality becomes very demanding is not an intrinsic feature of moral demands themselves but an extrinsic feature regarding the agent's social and political circumstances, (in Kant's example, an unfair and threating prince). Morality and happiness are more likely to come into conflict under very dire circumstances, such as political turmoil, instability, war and/ or oppression, when complying with everyday ordinary duties can become an almost impossible task (Pinheiro Walla 2022, 134). This means that our task is to create stable political and social conditions where human beings can fulfil their duties while pursuing their conception of happiness, a point that aligns well with my previous argument about the need to address the injustices in the social circumstances in our moral lives.

It turns out that on the reading of Kant that I have been developing in this article, Kantian ethics is much more palatable than is commonly thought, and this has allowed me to show that the Kantian position has some similarities to the Nietzschean position and answer some criticisms implicit in Owen's article. However, I have not been able to answer the two main problems associated with Nietzsche's second challenge. That is, I have not yet been able to fully explain why the pursuit of morality plus the pursuit of our overall pleasure should amount to a life of *self-affirmation* and ultimately *flourishing*, a task that I undertake in the next two sections.

VI. MORAL TRANSFORMATION AND SELF-AFFIRMATION IN KANT

In a recent article, Lucy Allais claims that in order to make sense of ourselves as agents, we have to interpret ourselves as basically good, well-ordered, unified selves acting under the principle of the moral law: "this is because Kant thinks that acting for reasons commits you to seeing what counts as a reason for action as governed by the constraint of respecting the humanity of others" (2021, 275). Moral transgressions, when they occur, constitute a serious threat to our self-understanding as unified, basically good, well-ordered agents. In order to avoid this threat, we are often lead to self-deception (we interpret ourselves as being better than our failures suggest) or despair (we realise that we are deeply flawed). Allais argues that since moral perfection is beyond our reach and since morality always represents a struggle for us, we will inevitably continue to transgress. These transgressions ultimately reveal that we do not have "a properly or fully ordered self" after all:

> The condition of virtue that is possible for us in the actual human condition involves not having a properly or fully ordered self. Human agents are not fixed, determinate characters, but rather messy, only partially unified works in progress…at best constantly engaged in a struggle to be better, including a struggle to be honest with themselves (2021, 276).

Given this lack of ordered character, Allais claims that forgiveness,[14] understood as a form of secularised grace, is needed as a response to each other's wrongdoing. Grace involves an overall view of the agent's character as "a possible (God's eye) perspective on our overall willing (our life seen as a whole) from which our flawed strivings can be seen as oriented toward good willing" (2021, 275) whereas forgiveness simply involves a view of an agent in relation to some specific wrongful action. Forgiveness involves evaluating agents "as oriented to better willing than their transgressions indicate" (2021, 277). Allais argues that we need to embark on this charitable interpretation of each other's wrongful actions to help us avoid despair and self-deception.

Allais's analysis is correct in pointing out that in order to make sense of ourselves as agents, we have to interpret ourselves as basically good, well-ordered, unified agents acting under the principle of the moral law. I also agree that moral development constitutes a struggle and that, therefore, transgressions are inevitable. However, I disagree with her claim that, in Kant, the actual human condition always involves not having a properly or fully ordered self and that it is therefore always possible to interpret a wrongdoer graciously or charitably as someone who is struggling towards the good . This interpretation of the human condition fails to account for the different forms of character and agency that, according to Kant's theory of radical evil, are implicit in different attitudes that we can adopt towards our own transgressions. To see this, we must appeal to Kant's theory of the three stages of radical evil: frailty, impurity, and depravity (Rel 6: 29-30). There are important differences between transgressions that occur 'after' the revolution of the heart (frailty and impurity) and the transgressions of those who are fundamentally committed to a maxim of self-love (depravity). In the case of depravity, as self-love is adopted as one's fundamental maxim, and thus as providing the ultimate justification for one's action, the agent will see her transgressions as fully justified fully justified and will be unlikely to condemn and repent them (or even recognise them as transgressions). In such cases, evaluating agents "as oriented to better willing than their transgressions indicate" may not only be mistaken but actually counterproductive counterproductive, since in the absence of genuine repentance, interpreting them as being oriented to better willing than their wrongdoing supports is not likely to help them to avoid despair and self-deception, and become better persons. In contrast, in cases of frailty and impurity, as Anastasia Berg has recently argued (2022), the subject adopts a fundamentally good maxim and thus their willing is oriented towards the good. However, although in these cases one's way of thinking is not corrupted, the problem is that in both cases the moral incentive is not sufficient determination to action. Frailty involves a form of weakness: we have the right maxim

14] The topic of Kantian forgiveness is a fascinating one and one that has often occupied me. My interpretation of Kantian forgiveness differs significantly from Allais' reading (see Satne 2016, 2018, 2020), but here I do not have space to enter into this debate.

but nevertheless perform a different (immoral) action. In the case of impurity**,** the agent performs the good action but only with the help of an incentive of inclination.

Frailty and impurity are, according to Berg's interpretation, *states of moral immaturity* when the agent has yet to achieve a stable moral character.[15] Here it is important to note that those frail agents who are still morally immature but who have undergone (or are undergoing) a revolution of the heart are likely to repent their transgressions and abandon their immoral maxims because of the perceived discrepancy between their actions and their principles. Although our motives are obscure so we cannot know in advance whether our repentance is sincere, the frail person has the right maxim so she is likely to feel remorse – the affective aspect of repentance- and attempt to reform her character, at least in the sense that she would attempt to strengthen her resolve to act on her morally good maxims. We can say that the nature of her commitment (to change) is revealed in the way that she keeps it and ultimately through her deeds. So, here we can say that her deeds are the criteria for her maxims. If she changes and starts to act differently because she has repented, then that is (defeasible) evidence of her moral transformation. Repentance must be reflected in a changed life and thus it provides (defeasible) evidence of a will oriented towards the good. Somehow paradoxically, repentance can have an affirming role. It is part of one's moral transformation which amounts to constituting oneself by affirming one's values.

Thus, contrary to Allais, acknowledging our moral transgressions, which for Kant would always involve guilt (CrP 5:98), is not always necessarily a burden because it can play an important role in our moral development. Guilt can also be analysed as having an intellectual and affective aspect. Although the affective aspect of guilt is painful (and, to that extent, a burden), the intellectual aspect of guilt, that is, the recognition of our transgressions, is bound up with the recognition of our freedom as the necessary condition of the very possibility of this recognition. As Alenka Zupančič stresses, in the Critical Elucidation of the Analytic of Pure Practical Reason, Kant shows "that which proves the reality of freedom — Or, more precisely, that which posits freedom 'as a kind of fact,' is presented…in the guise of guilt" (2000, 25). The key point for us is that recognising our transgressions as transgressions (the intellectual aspect of guilt) although painful (the affective side of guilt) leads us to recognise that we are free. As I put it in previous work:

> Guilt, in the intellectual sense, consists in passing a guilty verdict regarding our violation of the moral law. That is, when we reflect lucidly about our (objective) past wrongful acts, we cannot avoid reaching a guilty verdict. The mere fact that we judge ourselves as guilty shows that we take personal responsibility for our actions, and that, on the ultimate analysis, we think we could have acted otherwise in those same circumstances. This is evidence that there is an alternative description of actions, namely one appealing to a causality of freedom, which transcends the determinism of empirical causes (Satne 2021, 1515)

15] Berg 2022.

Thus, recognizing our guilt does not *necessarily* lead us to despair. It is through our guilt that we become aware of our fundamental freedom and this awareness represents an opportunity for change and improvement. This can also be seen as a (more optimistic)[16] way of acknowledging the significance of *I did it*. Perhaps through this process of acknowledging and repenting our immoral deeds, we can be hopeful (if not fully confident) about our prospects of becoming better people. This process of self-transformation is a process of self-affirmation and self-creation in which one constitutes oneself as self. In this process we affirm our values (our maxims as our commitments) and ultimately this involves *valuing oneself as the bearer of those values and commitments*.

In his Rousseau Lecture "On Vindication in Ethical Life", Owen distinguishes the ethical concepts of justification and vindication. Justification involves acknowledging the claims of what we owe to each other while vindication refers to the process of struggle to realize outcomes of value to the flourishing of the agent's life such that the agent has reason to affirm (or at least not regret) what they have done. We can say that there is an element of vindication in Kant's position. Living a moral and happy life is the agent's goal (with all the subgoals and challenges that this involves) and one is accountable to oneself for attaining this life. However, the priority of justification (over vindication) is not (rationally) optional for Kant. The Kantian agent has a responsibility to herself to live her life in a way that is accountable to others.[17] A life in which we accomplish our goals at the cost of disrespecting others cannot ultimately be vindicated (or affirmed).

VII. BY WAY OF CONCLUSION: FLOURISHING IN KANT

To the extent that this process of moral transformation is also a process of self-affirmation which involves valuing oneself as the bearer of values, we can say that a life committed to a project of moral transformation is a life of flourishing. However, this statement does not fully capture all aspects of Kant's position. As we have seen in section V, Kant argues that happiness cannot be identified with the contentment that follows from complying with morality. Although morality is the highest (in the sense of supreme) good, it is not the whole good. The highest (in the sense of complete) good for a human being "must be understood…as [the] synthetic connection between [the] two heterogeneous, irreducible components"[18] of virtue and happiness. As we saw, the moral agent is also committed to the pursuit of their own happiness (albeit not unconditionally). Although morality and happiness are not intrinsically incompatible, they are likely to conflict under unstable and unjust external conditions. Thus, the

16] Arguably Kant's more optimistic stance derives from his rejection of naturalism, which constitutes an important disagreement with the Nietzschean position.

17] Incidentally, Pinheiro Walla also demonstrates, against Williams, that Kant's ethics does not require strict impartiality. See her excellent chapter 4 (2022).

18] Pinheiro Walla 2022, 55.

extent to which we can live a completely morally good life is not independent of social and political conditions. Paradoxically, having morality as our supreme end allows us to see the natural and social evils that plague humanity, not as something that we have no hope to improve, but as presenting us with a moral task, that is, as something we have a duty to, and indeed can, address. Ultimately, in the Kantian picture, the task of moral self-improvement also has a social dimension: we must avoid being complicit with the injustices of our time and instead work towards overturning them. It is in fair (or at least fairer) social and political conditions that we can pursue both moral and personal ends and flourish.

P.Satne@leeds.ac.uk

REFERENCES

Allais, Lucy. 2015. What properly belongs to me: Kant on giving to beggars. *Journal of Moral Philosophy* (12): 754-71.

———. 2021. Frailty and forgiveness. In *Forgiveness and its moral dimensions*, edited by Warmke, B., Nelkin, D. K., and McKenna, M. Oxford University Press.

Berg, Anastasia 2022. Evil or only immature? Kant and the complexity of moral failure. In *Rethinking Kant Volume 6*, edited by Valdez, E. Cambridge Scholars Publishing.

Cherry, Myisha. 2021. *The case for rage: Why anger is essential to anti-racist struggle*. Oxford University Press.

Drogalis, Christina. 2013. Kant's change of heart: Radical evil and moral transformation. Dissertations. Paper 512. http://ecommons.luc.edu/luc_diss/512.

Guay, Robert. 2002. Nietzsche on Freedom. *European Journal of Philosophy* 10: 302-27.

Korsgaard, Christine. 1996. *Creating the kingdom of ends*. Cambridge: Cambridge University Press.

Kant, Immanuel. 1991[1797]. *The Metaphysics of Morals*, translated M.J. Gregor. Cambridge: Cambridge University Press.

———. 1997a[1774-1777]. Moral philosophy: Collins's lecture notes. In *Lectures on Ethics*, trans: Peter Heath and edited by Peter Heath and J.B. Schneewind. Cambridge: Cambridge University Press.

———. 1997b[1797-1798]. Kant on the metaphysics of morals: Vigilantius's lecture notes. In *Lectures on Ethics*, translated by Peter Heath and edited by Peter Heath and J.B. Schneewind. Cambridge: Cambridge University Press.

———. 1997c[1785]. *Groundwork of the Metaphysics of Morals*, translated by M. Gregor and introduction by C.M. Korsgaard. Cambridge: Cambridge University Press.

———. 1998[1793]. *Religion within the Boundaries of Mere Reason and Other Writings*, translated by A.W. Wood and G. di Giovanni and introduction by R. M. Adams. Cambridge: Cambridge University Press.

———. 2002[1788]. *Critique of Practical Reason*, translated by W.S. Pluhar and introduction by S. Engstrom. Indianapolis: Hackett Publishing Company, Inc.

Morgan, Seiriol. 2005. The missing formal proof of humanity's radical evil in Kant's religion. *The Philosophical Review*, 114(1): 63-114.

Nietzsche, Friedrich. 1996[1887]. *On the Genealogy of Morals*, translated by D. Smith. Oxford, Oxford University Press.

———. 1998[1889]. *Twilight of the Idols*, translated by. D. Large. Oxford, Oxford University Press.

———. 2001[1882]. *The Gay Science*, ed. B. Williams, trans. Josephine Nauckhoff and Adrian del Caro Cambridge. Cambridge University Press.

Owen, David. 2009. Autonomy, self-respect, and self-love: Nietzsche on ethical agency. In *Nietzsche on freedom and autonomy*, edited by Gemes, K., and May, S. Oxford University Press, USA.

———. 2023. On vindication in ethical life. *Public Reason* 14(2)-15(1): 3-18.

Pinheiro Walla, Alice. 2022. *Happiness in Kant's Practical Philosophy: Morality, Indirect Duties, and Welfare Rights*. Rowman & Littlefield.

Reath, Andrews. 2006. *Agency & autonomy in Kant's moral theory. Selected essays*. New York: Oxford Clarendon Press.

Ridley, Aaron. 2009. Nietzsche's intention: What the Sovereign Individual promises. In *Nietzsche on freedom and autonomy*, edited by Gemes, K. and May, S. Oxford University Press, USA.

Satne, Paula. 2013. Reliability of motivation and the moral value of actions. *Studia Kantiana* 14: 5-33.

———. 2016. Forgiveness and moral development. *Philosophia* 44(4): 1029-1055.

———. 2018. Forgiveness and punishment in Kant's moral system. In *Kant's Doctrine of Right in the 21st Century*, edited by L. Krasnoff, N. Sanchez Madrid, and P. Satne. Cardiff: University of Wales Press.

———. 2020. Kantian forgiveness: Fallibility, guilt and the need to become a better person: Reply to Blöser. *Philosophia* 48(5): 1997-2019.

———. 2021. Kantian guilt. In *The Court of Reason: Proceedings of the 13th International Kant Congress*, edited by Himmelmann, B. and Serck-Hanssen, C., 1511-1520. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110701357-149

Sussman, David. 2005. Perversity of the heart. *The Philosophical Review* 114(2): 153-77.

Williams, Bernard. 1993. *Shame and necessity*. Berkeley: University of California Press.

Williams, Garrath. 2018. The social creation of morality and complicity in collective harms: A Kantian account. *Journal of Applied Philosophy*. https://doi.org/10.1111/japp.12334

Wood, Allen. 1999. *Kant's ethical thought*. Cambridge: Cambridge University Press.

Zupančič, Alenka. 2000. *Ethics of the real. Kant, Lacan*. London and New York: Verso.

Public Reason 14(2)-15(1): 57-59

Kant's Account of Personal Development: Comments on Paula Satne

Jakub Szczepanski
Jagiellonian University

Abstract. In this paper, I focus on Paula Satne's discussion of moral development in her "Moral Development, Repentance, and Self-affirmation". I argue that, preceding moral development, we can identify in Kant a notion of existential development that foreshadows some of Nietzsche's ideas and emphasises the existential potential of the philosopher from Königsberg.

Keywords: moral development, existential development, Enlightenment, emancipation, courage, autonomy, Paula Satne, Kant, Horace, Nietzsche.

In reference to the article ["Moral Development, Repentance, and Self-Affirmation"], I would like to pose a question regarding the issue of development raised by the author. Specifically, I am interested in whether Kant's concept of development is exclusively moral in nature. I have serious doubts about whether the foundation of Kant's concept of development can be understood solely in this way. Why? I have in mind the famous essay: "An answer to the question: What is enlightenment?" In it, Kant urges us to have "resolution and courage"[1] to "use our understanding without guidance from others. Sapere aude! [dare to be wise] Have the courage to make use of your own understanding!"[2]

Of course, this is just the beginning of the process of human emancipation and breaking free from "minority". However, this initial step demonstrates the fundamental basis for all development. So, why is independent reasoning so important, and what does Kant really mean? It seems to me that it is about the significance of continual reflection on, among other things, our own actions. As human beings, we tend to delegate our thinking to others (the "guardians"). We prefer to rely on pre-established thinking patterns. We like to act under the direction of others. There is nothing wrong with this, as long as we embrace such patterns reflectively. Philosophically, it is also about beginning to think autonomously ("walk alone"). According to Kant, only such thinking can yield fully moral outcomes. The second, or rather the first in a certain hierarchy, even more intriguing element is the path to independent thinking. To commence thinking independently and autonomously, one needs "resolution and courage". Why does thinking require courage when a rational mind seems far removed from any emotions? Well, non-independent thinking, which relies on an external element, is comfortable and tends to become dogmatic, a way of thinking that we consistently and universally

1] I. Kant, An Answer to the Question: What is Enlightment? (8: 36).

2] Ibid.

deem as correct. In such a mode of thinking, we feel comfortable because it becomes our secure refuge. Everything is familiar and as it has always been; the behavioral pattern is well-known to us and functions as expected. Moreover, we do not fully sense responsibility for decisions; rather, we follow instructions. After all, someone else makes decisions for us. What happens when the threat of impending change emerges? We feel fear! Fear of the unknown. Fear of losing our safe haven. Fear of embarking on the journey alone. To think in a new and perhaps risky manner, in a way previously unknown to us, we require courage. Kant is absolutely correct, or rather, not Kant, but Horace, the Roman poet he quotes in his pivotal call, "sapere aude!" The maxim cited by Kant actually originates from a letter to Maximus Lollinus. In the entire fragment from which Kant borrowed only two words, we can read that "Well begun is half done",[3] which clarifies why courage and determination are so significant here. Everything starts with them, and those who begin have already accomplished half of the work. Dare to be wise! Horace presents a situation in which someone who merely waits is akin to a simpleton who encounters a river on his path and waits on the bank, hoping that the water will cease flowing so he can continue. However, the river will not cease flowing, of course. What is the message of this story from the perspective of the Roman poet? Our independent thinking is necessary when we confront an obstacle that impedes our further progress. But why was the simple man (orig. *rusticus*), a mere villager, surprised by the presence of the river on his path? Because, as the story suggests, he did not know his path. Since he did not know it, he exhibited courage by venturing into the unknown, "walking alone". Thus, this individual had already initiated the process. When we undertake a task, we usually do not know what obstacles we will encounter on our journey. Resolution and courage seem to emanate somewhat from ignorance and a lack of knowledge about what truly lies ahead. On the other hand, the decision to embark on the journey is, in a sense, evidence of faith in our abilities. In reality, resolution and courage reveal themselves when we venture "beyond the threshold of our own home," when we depart from our secure haven. If we encounter an obstacle on our path, we cannot wait for the problem to resolve itself. Faced with a changing reality, we must respond wisely and autonomously. This means, once again, the necessity to leave our safe shelter. Therefore, we require resolution and courage; otherwise, we will not cross the "river". We will not surmount the emerging difficulties. We will remain stagnant, awaiting problems to resolve themselves.

Expanding on the entire context of Kant's call for wisdom, we perceive that resolution and courage, prerequisites for thinking, constitute crucial components here. The tragedy of the villager in Horace's story seems to stem from his inability to overcome the obstacle he encountered. Perhaps he only waited momentarily until the river ceased flowing, and ultimately, nature compelled him to take action?

3] Horace, Epistoles, 1.2

Returning to the question of moral development, if I were to identify commonalities between Kant and Nietzsche (Paula Satne references both of these philosophers), I would see a certain heroism in the struggle with oneself. Heroism that commences with resolution and courage to venture into the unknown. This is something that precedes moral decisions in a way. Something I would label as existential. Of course, moral questions also arise, but somewhat later. The individual standing before the flowing river more closely resembles Nietzsche's depiction of the "last man". The risks he undertakes, the falls he experiences on the path to enlightenment, are more reminiscent of a struggle for survival. A struggle with oneself. Essentially, if one delves into Kant, a distinctive dramatic quality akin to Nietzsche can be discerned in him. I would perceive this as common ground between both philosophers. However, this is, of course, my perspective. I observe Nietzsche's figure from the standpoint characteristic of Karl Jaspers or Hannah Arendt, who regard this philosopher as a figure "on the road", in the full existential sense of the term. Incidentally, examining Kant from the perspective of Nietzsche, the disrupter of order, constitutes a highly intriguing approach, revealing the seldom emphasized existential potential of the philosopher from Königsberg.

j.szczepanski@iphils.uj.edu.pl

Public Reason 14(2)-15(1): 60-76

To What Extent Does Kant's Doctrine of the Highest Good Embody a Realist Orientation Towards Ethics?

Sasha Mudd
University of Southampton

Abstract. In recent years, an increasing body of literature explores the idea that Kant may be read as endorsing a kind of meta-ethical constitutivism. In this paper, I argue that this label only partly fits. Although, it sheds helpful light on certain core elements of his ethical project in the *Groundwork*, it fits much less well when we consider his teleological re-casting of moral consciousness in his doctrine of the Highest Good. The way I show this is by examining Kant's evolving response to what, following contemporary debates, I call "the problem of alienation'. I argue that Kant solves a first version of this problem through his doctrine of respect, while staying true to his constitutivist commitments, but that this leaves a more complex version of the problem untouched. This second form of alienation, which I call "practical alienation' is, on my reading, the price Kant thinks finite agents must pay in order to be moral and is the core problem he explores in the Antinomy of Practical Reason in the second *Critique*. Through his doctrine of the Highest Good, I argue, Kant shows himself to be concerned with what is required for finite, embodied rational agents to *sustain confidence* in morality, despite what it costs them in practical alienation. According to this analysis, the transition to moral religion Kant articulates through the Highest Good leaves him at odds with his earlier, constitutivist commitments, and reveals a persistent tension in his ethics between what may be achieved from within the moral standpoint, as this is defined in constitutivist terms, and important practical goods this leaves out. Thus, I contend, many of the core concerns that motivate the doctrine of the Highest Good express what we might, following David Owen, regard as a realist orientation in ethics.

Keywords: highest good, constitutivism, alienation, realist orientation in ethics.

In recent years, an increasing body of literature explores the idea that Kant may be read as endorsing a kind of meta-ethical constitutivism (Korsgaard 1996a, 2009; Katsafanas 2013; Sensen 2017; Bagnoli 2019; Schafer 2019; Tenenbaum 2019;). By "meta-ethical constitutivism' I refer to the broad meta-ethical strategy that is informed by the following two claims: first, that the fundamental ethical or moral norms to which we are subject in acting are derived from the nature of what we are essentially; and second, that being subject to these norms, or conforming to them, at least to some extent, is part of what constitutes us as the kind of thing that we are (Korsgaard 2019). Is this the correct way of understanding the meta-ethical commitments expressed in Kant's ethics? My answer to this question, which I develop in this paper, is an ambivalent one: in a nutshell, both yes and no. Although, as I will argue, the constitutivist label sheds helpful light on certain core elements of Kant's ethical project in the *Groundwork*, notwithstanding certain doubts and caveats, the label nevertheless fits much less well when we consider Kant's teleological re-casting of moral consciousness in his doctrine of the Highest Good. The way I shall try to show this is by examining how Kant's evolving response to what, following contemporary debates, I will call "the problem

of alienation', leads him in the second *Critique* to a shift away from his earlier, clearer constitutivist commitments, most naturally associated with *Groundwork*.

More specifically, I seek to show that although Kant's account of respect solves a first, simpler version of the alienation problem, it nevertheless leaves a deeper and subtler version of the alienation problem intact. I call this practical alienation, a state in which one fails to see how acting morally is consistent with the fulfillment of one's essential nature taken as a whole, including one's sensible nature. In recasting morality teleologically through the doctrine of the Highest Good, I shall argue, Kant shows himself to be concerned with what is required for us to *sustain confidence* in morality, despite what it costs us in practical alienation. The transition to moral religion expressed in the Highest Good, however, leaves Kant at odds with his earlier, constitutivist commitments. What this reveals, I argue, is a persistent tension in his ethics between what may be achieved from within the moral standpoint, defined in constitutivist terms, and important practical goods this seems to leave out, namely what we might call the good of non-alienation. I end by reflecting on how Kant's concern with this tension reveals deep sympathies with what, following Owen, we may regard as a "realist" orientation to ethics.

In this paper, I proceed as follows. In Section I, I begin by laying out what I take to be the general case in favor of reading Kant as a constitutivist. My aim is to highlight the general plausibility of reading him this way without entering too far into the details that divide interpreters who nevertheless agree that he employs a constitutivist meta-normative strategy of some kind or another. Once I have made the case, I turn in Section II to examining how Kant's ethics fares with respect to a problem that has been identified as threatening the viability of the constitutivist approach more generally, that of alienation. I seek to clarify a first sense in which alienation poses a problem for the finite rational agent on Kant's view and argue that his theory of respect may be understood as an attempt to solve it. In Section III, I then go on to show how Kant's account of respect nevertheless leaves a deeper and subtler version of the alienation problem intact, one with which he grapples in the Antinomy of Practical Reason and tries, and in some measure fails, to overcome in his doctrine of the Highest Good. I call this practical alienation, a state in which one fails to see how acting morally is consistent with the fulfillment of one's essential nature taken as a whole, including one's sensible nature. In recasting morality teleologically through the doctrine of the Highest Good, I shall argue, Kant shows himself to be concerned with what is required for us to *sustain confidence* in morality, despite what it costs us in practical alienation. The transition to moral religion expressed in the Highest Good, however, leaves Kant at odds with his earlier, constitutivist commitments. What this reveals, I argue, is a persistent tension in his ethics between what may be achieved from within the moral standpoint, defined in constitutivist terms, and important practical goods this seems to leave out, namely what we might call the good of non-alienation. In Section IV, I conclude by reflecting on how Kant's concern with this tension reveals deep sympathies with what, following Owen, we may regard as a "realist" orientation to ethics.

I. THE CASE FOR READING KANT AS A CONSTITUTIVIST

By meta-ethical constitutivism I refer to a broad range of views that affirm or are implicitly committed to two claims: first, that the fundamental moral norms to which we are subject in acting are derived from the nature of what we are essentially (Schafer 2019, 177), and second that being subject to these norms, and/or conforming to them at least to some extent, is what constitutes us as the kind of thing that we are (Korsgaard 2019). In light of this highly abstract formulation, the question then becomes how to characterize "what we are essentially'. What understanding or concept of ourselves must be true so that the moral norms in question can be derived? Neo-Kantians take themselves to be following Kant in arguing that it must be the concept of a rational agent (O'Neill 1989; Korsgaard 1996a, 2009; Sensen 2017; Bagnoli 2019). Unsurprisingly, when it comes to reading Kant himself as a constitutivist, the trend, with some notable exceptions,[1] has been to read him as seeking to ground moral norms (or normative facts – facts about what is valuable, what there is reason to do, etc.) in facts about what is constitutive of rational agency. In this vein, discussion of Kant as constitutivist has naturally focused on the *Groundwork*. For instance, Paul Katsafanas writes:

> Kant attempts to anchor universal normative claims in facts about agency. An outline of the Kantian argument would go something like this: we are committed to acting autonomously. Acting autonomously requires acting on a law or principle. The law cannot be hypothetical, i.e., tied to the realization of some goal or the satisfaction of some inclination, because the will would then be determined to action by something external to itself (i.e., an inclination or goal). Instead, the law must be categorical; it must be unconditionally valid. Kant states the content of this law as follows: "act only in accordance with that maxim through which you can at the same time will that it become a universal law' (G 4:421). He argues that this law – the Categorical Imperative – rules out certain actions, thereby yielding determinate constraints on permissible actions. So, Kant moves from a claim about agency – that we are autonomous – to a normative claim about what we have reason to do (i.e., act on maxims that are in accordance with the Categorical Imperative) (2013, 35).

This move from some feature of the will to a normative standard from which normative reasons can be derived has been called "the basic constitutive move' (Ferrero 2018, 118). Although, of course, the force, meaning and success of Kant's arguments for the moral law as the supreme principle of morality in *Groundwork* are tremendously controversial, the basic structure of the arguments in *G* I and II – as here glossed by Katsafanas – is much less so. And while the best way of characterizing the feature of the rational will from which the moral law is derived may be subject to dispute[2], that Kant

1] According to Schafer, Kant's constitutivism is better read as a "reason-first' rather than an "agent-first' (2019, 179) constitutivism. The idea here is that Kant does not take the rational *agent* as the constitutive basis of norms, but rather *rational activity* just as such.

2] For related discussions, see O'Neill 1989, ch. 7, Korsgaard 1996b, ch.3, Wood 1999, ch.3, Guyer

intends to derive the law in this way in the *Groundwork* is not on the whole controversial, and even less so for those who read him as constitutivist (Korsgaard 1996a, 2009; Bagnoli 2019; Schafer 2019; Tenenbaum 2019; Sensen 2017; Katsafanas 2013).

For the sake of advancing my core argument, I will not say more about this here, except just to note that for Kant the fact that the moral law is derived from the nature of the rational will taken alone is, of course, captured by him in the very idea of autonomy. In the *Groundwork*, he defines "autonomy of the will [des Willens]' as "the property of the will by which it is a law to itself (independently of any property of the objects of volition)' (G 4: 440, cf. 4: 447). He goes on to say that autonomy of the will is "the supreme principle of morality' (G 4: 440). In the first instance, this means that the principle of morality is itself autonomous in the sense that it is derived from the nature of the rational will alone and from no other source.

Despite consensus on these very broad points, however, reading Kant as a constitutivist in this way throws up doubts and difficulties,[3] which are not easy to resolve without further detailed analysis. I shall flag one such issue in particular, for it is particularly relevant to the question of whether and to what extent Kant remains true to his apparent constitutivist leanings even after having recast moral consciousness through his doctrine of the Highest Good.

Kant, like all constitutivists, may be read as arguing that some feature of the will or of the capacity for rational agency both constitutes events as actions and generates a standard of assessment for action, from which standard normative claims are then derived (Katsafanas 2013, 35). Thus, like any constitutivist, Kant must solve the so-called "bad action problem' (61-62), that is, he must open a gap between action that possesses this constitutive feature and action that satisfies the success criterion derived from it . Opening this space is tantamount to offering an account of bad action, i.e. action that counts as such in virtue of possessing the constitutive feature, while yet failing to satisfy the success criterion derived from it. In the absence of such a gap, all action would count as good by default and the conceptual space required to identify defective action would be foreclosed. We can frame the problem in Kantian terms by asking: If autonomy is constitutive of the rational will, how can it be that this feature of the will generates a normative requirement to act in accordance with a principle of autonomy? Was not autonomy just supposed to have been constitutive of rational action as such?

As is well known but, I think, underthematized by those who read Kant as a constitutivist, he responds to this problem by distinguishing between the holy will, whose acts are indeed constitutively determined by the moral law, and the finite will, whose acts ought to be constitutively determined by that law, but may not be. Many of us will be familiar with the rich passage we find at *G* 4: 413, where Kant explains why imperatives do not hold for the divine or holy will. He writes:

2007, ch.5, Reath 2006, ch.7, Nyholm 2015, ch. 2. Kleingeld 2017.

3] See, for example, Street 2012.

> [...]The ought is out of place here, because holy volition is of itself necessarily in accord with the law. Therefore, imperatives are only formulae expressing the relation of objective law of volition in general to the subjective imperfection of the will of this or that rational being, for example, the human will (G 4: 413-14).

In a nutshell, because the holy will is perfectly rational, the laws of rational volition as such determine its activity without any kind of normative or – to use Kant's word – "necessitating' force. Such laws are merely descriptive of its necessary form of activity. Whereas because the same laws have necessitating force for us, they are normative, and we may either succeed or fail to act in accordance with them. It follows then that claiming that the principle of autonomy is constitutive of rational agency as such is ambiguous. What is constitutive of *finite* rational agency is the normative demand to live up to this principle.

Appreciating Kant´s way of solving the so-called bad action problem compels us to refine our description of his constitutive approach to normative justification outlined earlier. We stipulated above that constitutivists affirm that some feature of the will (or of rational agency) both constitutes events as actions and generates a standard of assessment for action, from which standard normative claims are then derived. Kant, in turn, was presented as grounding moral norms in facts about what is constitutive of rational agency. But the way in which Kant deploys the distinction between the holy will and the finite will complicates this story, showing it to be misleading if not downright incorrect.

As Kant's explanation of necessitating force makes plain, it is not the case that he derives a norm of conduct from the nature of the rational will considered alone, without reference to anything else. For he appeals to the sensible nature of imperfectly rational, finite beings like us in order to account for the imperatival form the moral law takes in relation to us. Normativity, for Kant, is thus grounded in the subjective presentation of an objective law and is therefore a relational property of the law as it is cognized by a finite, sensibly conditioned rational beings. Strictly speaking, then, the nature of practical reason alone does not determine the norms of its proper exercise, for taken alone it determines no norms at all. Instead, it determines the laws that descriptively characterize the activity of perfectly rational beings, like the holy will. As we have seen, in order to derive from these objective laws a norm that applies to us we must go beyond the nature of pure practical reason as such and appeal to the conditions of our sensible embodiment. For this reason, it might be argued that the kind of constitutivism that emerges from Kant's discussion is not best thought of as a variety of agential constitutivism pure and simple but instead as taking as its starting point the notion of the finite rational being with an essentially bifurcated nature: both rational and sensible.

Now this might seem like a relatively minor point, and one which hardly ought to tarnish Kant's constitutivist credentials, but I think it begins to shed light on other important dimensions of his ethics which also fit less well with the constitutivist label,

at least as it tends to get used in contemporary discussion. In order to see this, let us turn to Kant's way of dealing with the so-called alienation problem, which has recently been identified as posing a challenge for any constitutivist approach to normative justification.

II. KANT SOLVES THE PROBLEM OF MORAL ALIENATION THROUGH HIS DOCTRINE OF RESPECT

Rahel Jaeggi's influential book by the same title provides a useful point of departure. She writes:

> Alienation means indifference and internal division, but also powerlessness and relationlessness with respect to oneself and to a world experienced as indifferent and alien. Alienation is the inability to establish a relation to other human beings, to things, to social institutions and thereby also – ultimately – to oneself. An alienated world presents itself to individuals as insignificant and meaningless, as rigidified or impoverished, as a world that is not one's own, which is to say, a world in which one is not "at home' and over which one can have no influence through one´s actions. The alienated subject becomes a stranger to itself; it no longer experiences itself as an "actively effective subject' but rather as a "passive object' at the mercy of unknown forces (Jaeggi 2014, 3).

As this rich sketch illustrates, alienation has been taken to refer to a varied collection of maladies, which reveal a concern both with the subject's internal relation to herself as well as her relation to the world, while at the same time seeming to suggest that certain defective modes of self-relation lead to or imply defective modes of world-relation, and also – perhaps – vice versa.[4] It is notable that these mutually implicating defective modes of relation appear to hinge in some way on our incapacity to represent ourselves as efficacious agents, that is, as beings capable of making a difference to the way things are through our actions. While it is not my purpose to make headway in unifying the diverse phenomena that Jaeggi evokes in this passage, what I will do is connect these to what I regard as a widespread concern about the viability of constitutivist meta-ethical approaches. In a recent paper, Sergio Tenenbaum voices this concern in a particularly sharp way, arguing that a certain sort of alienation poses a challenge to constitutivists who seek to ground moral norms in facts about agency (Tenenbaum 2019). He writes:

> [...] Vulnerable to an important worry; namely, that it leaves us alienated from the moral norms that it claims we must follow [...] in a nutshell, it seems that constitutivism cannot provide an adequate account of the relation between the constitutive norms of agency and the particular ends that agents pursue [...]. even if constitutivism could show that norms of agency are inescapable – one of its chief aims – it would nevertheless leave us alienated from these norms. (163-64)

4] For other approaches to the notion of alienation, see Wood 2004, and Schmitt 2003.

The key thought here is that the constitutive end of agency is that end to which we find ourselves implicitly committed insofar as we act at all. In this sense, the norms derived from it abstract from the particular interests, inclinations and concerns that motivate the pursuit of all our other, particular or subordinate ends. But at the same time, the constitutivist must allow that the agent's implicit commitment to the constitutive end (and to following the norms derived from it) can, in principle, conflict with her commitment to the other ends she pursues. After all, if this were not the case, it would be difficult to conceive of the norms derived from the constitutive end as constraining the agents' choice of subordinate ends. The problem is that our capacity to care about and feel motivated by the constitutive end of agency is threatened by the very fact that this end seems to bypass the interests and inclinations that motivate the adoption of all our other ends. And yet, at least in principle, we need to be motivated by this end sufficiently enough for it to function as the constraint it is meant to be on our other choices. To meet this challenge, then, the constitutivist must show how our implicit commitment to the constitutive end of agency can become explicit and sufficiently motivating from the first-person point of view. For without such an account, it would be difficult to see why or how we could come to care about following the norms that are grounded in the constitutive end, especially when doing so conflicts with the pursuit of our subordinate ends.

Alienation, on this telling, refers to the agent's failure to stand in a correct affective or motivational relation to the ends and norms to which he is, supposedly, implicitly committed just in virtue of acting at all. To be alienated from the constitutive end of agency, is just, in the first instance, for the agent to fail to be motivated by this end sufficiently enough for it to function as the felt constraint it is meant to be on her choice of other ends and maxims.[5] So described, alienation clearly captures a broad concern raised, in one way or another, by many critics of constitutivism, often under other lables, a concern that echoes Enoch's famed Shmagency objection, which can be read as demanding an account of why an agent cannot simply choose to reject those norms that are putatively constitutive of agency on the grounds that she does not care to be an agent in the first place.[6]

With this general sketch of the threat alienation poses to constitutivists in hand, we are now in a position to translate this concern into a language that more closely tracks Kant's putative constitutivism. Alienation, on this re-telling, refers to the inner state

5] It is worth noting that at the highest level of abstraction, constituting oneself as an agent, by following the constitutive norms of agency whatever these are taken to be, can be construed as the constitutive end of agency. This is just to explain why one can move rather freely, I think, between talking of constitutive ends and constitutive principles or norms of agency.

6] It is worth noting that this problem holds even for those who reject the notion that agency has a constitutive end akin to the constitutive ends of games like chess or baseball. For, at the highest level of abstraction, constituting oneself as an agent, by following the constitutive norms of agency whatever these are taken to be, can be construed as the constitutive end of agency, as Enoch's shmagency objection plainly illustrates.

of the agent who fails to care about or be properly motivated by the ends of morality, here construed as constitutive ends. We might even say that alienation describes the experience of heteronomy from inside the first-person point of view, insofar as this latter state or principle of choice can be said to manifest our lack of respect, understood as the unique moral motive. For the heteronomous agent experiences the normative demands of pure practical reason as alien, or – in other words – she is alienated from them, because she has made the principle of self-love her guiding maxim. Thus, although constitutive ends reflect the pure rational part of the agent's essential nature, they can nevertheless still be experienced by her as alien or as alienating insofar as she lacks the motive needed to choose them, thereby making them "her own" in this richer sense. So, to be alienated, on this more explicitly Kantian analysis, is not merely to stand in an estranged, distant or faulty relation to the pure rational part of one's nature but to do so precisely insofar as one fails to choose or act on the ends constitutive of that part of one's nature.

It follows that the problem of alienation, which I shall henceforth dub "moral alienation', can be solved by Kant only if his moral psychology can explain how autonomy rather than heteronomy can motivate the finite rational agent. I submit that Kant was, in his own way, importantly concerned with this problem, and that his account of respect is meant to solve it (G 4: 401n; CPrR 5: 73-81). Due to constraints of space, I will not present my full argument here, but the important upshot for present purposes is just that through his doctrine of respect Kant takes himself to have shown *that we are in fact capable of taking a pure, a priori interest in morality*, and on the basis of this interest committing ourselves explicitly to our constitutive moral ends, whatever our countervailing empirical desires and interests may be. The motive of respect just is this unique a priori motive or interest, without which compliance with the constitutive norms of agency would be impossible for beings like us.[7]

However, on my view, Kant was far more worried about a different, more complex phenomena that may also be considered a species of alienation. In this case, as I shall argue, Kant was not able to find a solution to this more complex form of alienation from within the bounds of his constitutivism. In a nutshell, Kant´s doctrine of respect, read as a solution to the problem of moral alienation, leaves untouched a deeper and more intractable kind of alienation, one that does not concern how creatures like us can come to take a first-personal interest in acting morally, but instead concerns how we may come to view moral action as purposive for and consistent with the sensible part of our nature, and thus with the fulfillment of our practical nature taken as whole. This form of alienation, which I shall dub "practical alienation', concerns our rational need to see moral action as purposive in fulfilling the whole of who we are, as essentially bifurcated beings who are both sensible and rational. In other words, it concerns our need to see

7] For discussion on Kant doctrine of respect, see Herman 1981; Ameriks 1987; Allison 1990; Guyer 2000, Reath 2006, Singleton 2007, Noller 2019.

morality as compatible with what we might call human flourishing. In attempting to come to grips with the problem of practical alienation via his Doctrine of the Highest Good, I shall argue, Kant ends up betraying certain core features of his constituivist commitments.

III. KANT'S DOCTRINE OF THE HIGHEST GOOD AND THE PROBLEM OF PRACTICAL ALIENATION

So, what do I have in mind with the problem of practical alienation? The problem, in essence, is that just as heteronomous action motivated from the principle of self-love can leave us morally alienated from the pure rational part of our nature, so too moral action from respect, for the sake of pure rational ends, can leave us alienated from the empirical part of nature. Here I use "alienated" in the broadest sense to refer to a state in which the subject is motivationally or affectively estranged from some aspect of her essential nature where this estrangement is manifest in her incapacity to choose or act on the ends associated with that part of her nature in a full and complete way. In other words, although moral action from respect may be viewed as an expressive of who we really are (and freely choose to be) as rational beings, and thus as expressive of a non-alienated relation to the pure rational part of ourselves, that very same action may nevertheless leave us alienated from the animal part of ourselves, which is necessarily and inescapably interested in the heterogenous end of happiness (G 4: 415; CPrR 5: 25; R 6: 387). This can happen when acting morally leaves us with an unfulfilled yet still rationally legitimate need to see our permissible empirical ends realized and in harmony with what morality demands of us. It is this more complex manifestation of practical alienation that is, of course, Kant's central concern in the Antinomy of Practical Reason. In this sense, the possibility of the subject's standing in an alienated relation to herself is not exhausted by her possible estrangement from morality (for which Kant's doctrine of respect supplies a kind of solution), but also arises necessarily from the very nature of moral action itself.

The point, in a nutshell, is just that insofar as the ends of happiness and morality are fundamentally heterogenous and insofar as both are equally expressive of one or the other part of our essential nature, neither moral action nor action from the principle of self-love can be considered expressive of or aligned with our true nature taken as a whole. In this sense, it seems that there is no course of action or maxim of action that can leave us fully un-alienated, for just as sure as heteronomy expresses moral alienation, so moral autonomy threatens practical alienation.

Despite these structural symmetries, however, it is worth noting the ways in which these two different types of alienation differ from one another. In the first case, we can experience the ends of our pure rational nature as alien (in the moral sense) insofar as we fail to care about or be properly motivated by them, as is manifest in heteronomy. But, in the second case, the problem is not that we lack the motive or interest needed

to choose the ends of happiness. In fact, on the contrary, Kant says we simply cannot help but have and care about these ends. And yet, in acting morally, we cannot make the ends of happiness fully our own by incorporating them into our maxims in the usual sense, since, in acting morally, we are not aiming to realize these ends at all. This, on my reading, is precisely what happens with the permissible ends of happiness when we act from duty, as Kant suggests in the Antinomy. So, the idea is just that the ends of happiness, if left chronically denied and without hope of fulfillment, become alien to us in the practical sense of the term when we cannot see ourselves realizing them through our moral action. Thus, for Kant, the solution to moral alienation, action from the motive of respect, in fact creates the conditions of practical alienation, insofar as respect abstracts completely from all ends and motives associated with happiness.

Due to constraints of space, I will not enter into the exegetical arguments that support this reading of the Antinomy. Instead, I will turn to the general conclusions I seek to draw from my analysis, which concern Kant's attempt to solve the problem of practical alienation through his doctrine of the Highest Good, and the ways in which this pulls against his earlier constitutive commitments.

So, in what sense does Kant intend his doctrine of the Highest Good to offer a a solution to the problem of practical alienation? As is well known, Kant argues that there is no other way of resolving the conflict between the demand for virtue and our need for happiness except through thinking the law of morality in teleological terms as commanding the realization of the Highest Good (CPrR 5: 108-14), where this is a state in which perfect virtue *has brought about* an exact proportion of happiness.[8] On my reading, then, the Highest Good articulates a necessary ideal of *non-alienated moral action* projected as a moral world in which our rational and sensible natures have been harmoniously integrated and reconciled. This ideal allows us to look at morality as though it were purposive for the whole of who we are, as bifurcated beings who are both sensible and rational, and, thus, as purposive for our happiness too. The harmonious fulfillment of the complete set of ends associated with our nature taken as a whole, including the permissible ends of happiness, is as close as Kant can get to providing a general vision of human flourishing that is consistent with morality.

Importantly, however, the conditions under which the Highest Good may be thought possible as is necessary for the resolution of the antinomy – are immortality of the soul, so that we may hope for the complete realization of virtue, and the existence of God (CPrR 5: 124), so that we may hope that virtue may yet be capable of causing happiness, which causal relation is nowhere to be observed in empirical nature.[9]

8] The argument is in fact slightly more complex than I have presented it. Kant casts the Highest Good as the necessary object of a moral will, and therefore as commanded necessarily by the moral law. He reasons that if the possibility of the Highest Good generates contradiction, this would prove the moral law invalid, thereby producing *a reductio ad absurdum*, given that the validity of the law has been established (CPrR 5:111-14).

9] In fact it is not wholly clear on the face of it whether Kant thinks virtue must be conceived as

Achieving the highest good is thus beyond our agential control, requiring God (and the afterlife) for its fulfillment. Thus, in projecting the teleological fulfillment of moral action in onto an ideal plane, Kant makes it unachievable for sensible creatures who can, at best, exercise control over the selection of their own maxims, without thereby ensuring anything about the sort of world they thereby help to bring about, including how much happiness that world might contain and in what distribution. In effect, then, the Highest Good offers us a vision of what perfect, un-alienated rational volition would look like, if it were possible, for beings like us who are, as a matter of fact, inescapably alienated in the empirical world we know due to our finite, bifurcated nature.

This, however, seems to raise a problem for Kant´s constitutivism. For if it turns out that the Highest Good is the constitutive aim of rational agency, (that is, if the CI can be formulated as a command to seek the Highest Good, as Kant states that it can be), then there appears to be no clear moral success criterion for action. The problem can be expressed in a question: How can it be the case that in exercising our rational agency we are constitutively committed to realizing the Highest Good, if this necessarily involves achieving a certain relation to nature that lies beyond our agential control? A success criterion for action must specify actions that we are, in principle, capable of carrying out. Put another way, if the Highest Good were the constitutive end of agency, then all our actions would count as bad or defective, making the criterion in question no criterion at all.

Now, it may be objected that this problem is not genuine but only apparent, because what Kant really means is that the CI can be reformulated as a demand to *seek* the Highest Good (not to realize it), where this seeking is perfectly consistent with the end in question remaining unrealizable for agents like us. But this cannot be right, since reason can no more consistently command us to *seek* to attain the unattainable as it can command us to attain the unattainable. Of course, seeking to attain the unattainable is something we *can* in fact do (tragically, it is probably something we do all the time), but *reason* cannot consistently ask this of us. For, as Kant repeatedly insists, an end can only be an end for us in so far as we take it to be realizable. This is a requirement of rational consistency. If it were not, Kant would have no *rational* grounds for justifying our acceptance of the practical postulates. For, recall that their rationally justified function is precisely to allow us to see the Highest Good as something that, with the help of God, we may yet attain in the afterlife. By the same token, Kant´s argument that we do have rational warrant to accept the postulates is an admission that without God and the afterlife, the Highest Good is not attainable for us and we have no rational warrant to see it otherwise (in this life, with respect to our own powers). Insofar, then, as we judge the attainability of our ends with respect to this world, and with respect to our own agential

---

bringing about happiness according to the laws of nature in this life, or in the afterlife, where the laws of nature presumably no longer apply. I remain neutral on this question, which need not be resolved for present purposes, although in what follows I refer to God (rather than to God and immortality both) as shorthand for the practical postulates.

control, a command to *seek* that which we know to be unattainable simply cannot be rational. In sum, the ought-implies-can principle, as applied to practical life in this world, means that the CI, interpreted as a command to seek the Highest Good, is no success criterion for action at all. If, however, we abandon this problematic teleological formulation of the moral law and concede that the constitutive norm governing action is limited to the CI as it is formulated in *Groundwork,* then we are, once again, left with a version of the practical alienation problem Kant identifies in the Antinomy.

My conclusion, then, is that there is a persistent tension in Kant's ethics between what we might regard as "the morality system" and what it leaves out, one which cannot be solved from inside ethics alone, but requires a transition to religion, that is, requires belief in the practical postulates and all they bring with them. This is because Kant sees rational animals in the empirical world as essentially and unavoidably prone to alienation in their activity. In other words, precisely insofar as Kant's ethics is in the business of providing us with a success criterion for action based on the constitutive features of pure practical reason, it cannot be in the business of solving our practical alienation problem, that is, it cannot be in the business of providing us with a comprehensive account of human flourishing, at least not from inside the ethical perspective. Insofar as it does try to be in this business, it fails at its first task, namely that of generating a standard of right conduct from autonomy understood as the constitutive aim of rational agency. In conclusion, no answer to the problem of practical alienation can be found within Kant´s moral framework, insofar as this is interpreted in constitutivist terms, for the structure of finite rational agency is precisely that which makes moral action – in this world – necessarily alienating for us in the first place. The link between doing one´s duty and flourishing, understood as the fulfillment of our nature taken as a whole, cannot be secured from inside the ethical standpoint without resorting to God and religion.

IV. DOES KANT'S DOCTRINE OF THE HIGHEST GOOD EVINCE A REALIST ORIENTATION IN ETHICS?

David Owen, in his chapter entitled "Realism in ethics and politics: Bernard Williams, political theory and the critique of morality', explores the meaning of Williams' realist orientation in ethics, and in thinking more broadly (2018). As one might expect, Kant largely emerges as the prototypical baddy in this analysis. But although Kant is indeed guilty of some of the charges brought, I submit that he can nevertheless be cleared of others with relative ease. This is because his reworking of morality through the Highest Good, as I have presented it above, in fact exemplifies some of the central tenets of the realist orientation. This result sheds unexpected light on the heterogeneous ways in which a broadly realist approach to ethics can be articulated and helps underscore the important ways in which the ethics of the second *Critique* depart from many of the core commitments that predominate in the *Groundwork*.

Let me begin by noting the ways in which Kant is clearly *not* in sympathy with the realist. First, he clearly rejects realism´s methodological principle that says, "never explain the ethical in terms of something special to ethics if you can explain it in terms that apply to the non-ethical as well" (Williams 1995b, 204). Owen is right to characterize Kant as embracing the opposite strategy. While the realist tries to understand human moral capacities in terms of psychological materials we use anyway elsewhere, Kant goes the opposite route, positing a unique faculty, pure practical reason, by which specifically moral truths are apprehended, and the possibility of moral motivation is accounted for. In this respect Kant is clearly out of step with the realist's central methodological orientation.

This makes it all the more surprising, then, how closely Kant aligns with the realist in at least two other respects. First, on my reading, Kant largely shares the realist's "pessimism of strength", a phrase coined by Nietzsche and adopted by Williams in *Ethics and the Limits of Philosophy* to characterize his realist outlook. Second, with the realist, Kant rejects the thought that the foundational justification of the ethical outlook requires us to reconcile outside and inside perspectives on ethics. In what follows I take each of these points in turn.

The pessimism of strength associated with the realist orientation to ethics, as Owen reads it, is intended as an alternative to the optimism of the philosophical tradition. Owen quotes Raymond Guess at length in describing the five key features of the optimism to which the pessimism in question stands opposed. In going one by one through this list, we shall see just how closely Kant aligns with the pessimist against the optimist, despite what many seem to assume.

"First of all, traditional philosophers assumed that the world could be made cognitively accessible to us without remainder". (Geuss 2005, 223) Here it should go without saying that Kant assumes no such thing. On the contrary, he holds that all we can know are appearances, never things in themselves. "Second, traditional philosophers assumed that when the world was correctly understood, it would make moral sense to us." (223) Here again, Kant thinks just the opposite. Understanding the world brings us not one inch closer to understanding morality, for Kant, because freedom does not belong to the world of appearance at all. For Kant the world that is cognitively accessible to us makes no moral sense whatsoever, because the laws of nature and the laws of freedom refer to independent domains of legality and therewith intelligibility. Geuss continues, "Third: the kind of "moral sense' which the world makes to us is one that shows it to have some orientation towards the satisfaction of some basic, rational human desires or interests, that is, the world is not sheerly indifferent to or perversely frustrating of human happiness." (223) On the reading of the second *Critique* I have advanced in this paper, Kant rejects this thesis completely. For him, the empirical world is in fact perversely frustrating of human happiness insofar as it offers no evidence whatsoever that virtuous conduct is correlated with happiness in the empirical world. It is precisely this total lack of coordination between virtue and happiness that requires us to look beyond the world – to God and the afterlife

– in order to conceive how their mutual coordination might yet be possible. So far, then, Kant is with the pessimist, against the optimist, point for point.

Let us continue with Geuss' list. "Fourth, the world is set up so that for us to accumulate knowledge and use reason as vigorously as possible will be good for us and will contribute to making us happy." (223) Here, Kant's position is more complicated, and much will depend on exactly what is meant by the vigorous use of reason being "good for us", but on the whole I think it is fair to say that Kant rejects much of what is implied in this statement. The world is precisely not set up to ensure, guarantee or even make likely that the vigorous use of reason will contribute to making us happy. Again, this is why Kant thinks we need the practical postulates in order to imagine the conditions under which rational, moral effort might reliably bring about happiness in proportion to virtue. It is true, however, that Kant thinks that the accumulation of specifically moral knowledge is "good for us," but "good" here functions in a special, limited sense, bearing no direct relation to happiness.

"Finally", writes Guess, "these traditional optimist philosophers assumed that there was a natural fit between the exercise of reason, the conditions of healthy human development, the demands of individuals for satisfaction of their needs, interests and basic desires, and human sociability." (2005, 223). Again, for Kant, as I´ve been arguing, the opposite is in fact the case. In sum, then, Kant appears to be much more closely aligned with the pessimism of the realist than with the optimism of the philosophical tradition, as Owen, following Geuss, describes it.

There is, however, a second surprising way in which Kant turns out to agree with central tenets of the realist orientation, despite what Owen occasionally seems to suggest. On my reading, Kant goes along with Owen's realist in rejecting the thought that the foundational justification of the ethical outlook requires us to reconcile outside and inside perspectives on ethics. In order to grasp this, it will behoove us to look more closely at what Owen says about inside and outside perspectives. He writes: "[...]One way to think about the grip of "morality' can be elucidated by noting the predicament that arises once it becomes impossible truthfully to hold on to Aristotle's natural teleology as a way of justifying ethical life" (81). Here Owen proceeds to quote Williams at length.

> The question of the justification of ethical life can be asked from a perspective that is "inside' ethical life or a perspective that is "outside' ethical life, where the former asks what reasons we have for continuing to live such a life, where the first personal reasons invoked may draw on the ethical dispositions that agent takes to be part of who they are, and the latter asks why we should take up ethical life at all, where the third personal reasons invoked take those ethical dispositions as objects of evaluation. For Aristotle, the virtuous agent experiences no conflict or tension between inside and outside perspectives because, on Aristotle's theory, there is a view of "a certain kind of ethical, cultural and indeed political life as a harmonious culmination of human potentialities, recoverable from an absolute understanding of nature. (Williams 2011, 59)

Williams central claim here, on Owen's telling, is that once the Aristotelian assumptions which fitted together the agent's perspective and the outside view have collapsed, then the justification of ethics becomes much more problematic. Once this gap has been opened, Williams writes,

> "[w]e understand – and, most important, the agent can come to understand – that the agent's perspective is only one of many that are compatible with human nature, all open to various conflicts within themselves and with other cultural aims. With that gap opened, the claim I expressed by saying that agent's dispositions are the 'ultimate supports' of ethical value takes on a more skeptical tone. It no longer sounds enough." (Williams 2011, 59)

Owen then maps out a potential response to this predicament, according to which one can "surrender the project of giving an account of a fully developed human life and [instead] adopt a way of justifying ethical life in terms of 'morality' through an appeal to rational agency in which "morality' presents itself to the rational agent as a categorical demand. On this view, there is no tension between an inside and outside perspective since the reasons to be moral are intrinsic to rational agency as such." (Owen 2018, 82)

Here it is clearly the constitutivist Kant that Owen has in mind. But as I have argued in this paper, the tension between what we might think of as an inside and an outside perspective on ethics remains within Kant´s system, since it is precisely this tension that, on my reading, Kant's doctrine of the Highest Good seeks, however unsuccessfully, to resolve. That is to say, contrary to Owen, it is not the case that for Kant there is no tension between an inside and outside perspective insofar as our reasons to be moral are constitutive of or intrinsic to rational agency as such. Such a tension would plausibly disappear if we were solely pure rational beings, but instead, for Kant, we are essentially bifurcated beings, both sensible and rational. It follows that the moral norms constitutively derived from the pure rational part of our nature do not prevent us from asking justificatory questions from a different normative standpoint, namely that defined by our animal nature, with its distinctive set of ends and interests. On Kant's account, then, we are rationally compelled to step "outside" ethics in order to ask about the fate of happiness, construed as that end to which we are inescapably attached though a kind of "natural necessity".

But, importantly, for Kant, the project of attempting to reconcile our need for happiness with the demands of morality through postulating the Highest Good as the end of pure practical reason is not undertaken for the sake of justifying morality. This is not Kant's view at all. Rather, morality and its chief law are fully justified without any appeal to ends, even that of the Highest Good. Instead, in establishing the Highest Good as the end or pure practical reason Kant is trying to re-orient us, practically, to the task that morality sets us, which is justified on independent grounds. This project of practical reorientation, in turn, displays a concern that Owen takes to characterize the realist orientation, namely a concern with our ability to sustain confidence in the moral outlook, whatever its justificatory supports. So, in this final sense, then, Kant sides with

the realist in rejecting the thought that foundational justification of the ethical outlook requires the reconciliation of outside and inside perspectives. Kant wants to reconcile these perspectives, to be sure, but with the sole purpose of helping us sustain rational confidence in the moral outlook, where this involves offering us hope of escaping the chronic state of alienation which characterizes practical life for us here on earth. Seeing how Kant, despite his anti-naturalism, still retains key elements of moral realism serves to expand our notion of the possibilities afforded to us through this basic orientation.

alexandra.mudd@uc.cl

REFERENCES

Allison, H. 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.

Ameriks, K. 1987. The Hegelian Critique of Kantian Morality. In *New Essays on Kant*, edited by B. Den Ouden and M. Moen, 179-212. New York: Peter Lang.

———. 2003. *Interpreting Kant's Critiques.* Oxford: Oxford University Press.

Bagnoli, C. 2019. Authority as a Contingency Plan. *Philosophical Explorations*22(2): 130-45.

Ferrero, l. 2018. Inescapability Revisited. *Manuscrito – Rev. Int. Fil. Campinas* 41(4): 113-58.

Guyer, P. 2007. *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide*. London: Continuum.

———. *Kant on Freedom, Law, and Happiness.* Cambridge: Cambridge University Press.

Herman, B. 1981. On the Value of the Motive of Acting from Duty. *Philosophical Review* 90: 359-82.

Jaeggi, R. 2015. *Alienation*. New York: Columbia University Press.

Kant, I. 1991. *The Metaphysic of Morals.* Trans. M. Gregor. Cambridge: Cambridge University Press.

———. 1997. *Groundwork of Metaphysic of Morals.* Trans. M. Gregor. Cambridge: Cambridge University Press.

———. 1998. *Religion within the Boundaries of Mere Reason.* Translated by A. Wood and G. di Giovanni. Cambridge: Cambridge University Press.

———. 2012. *Lectures on Metaphysics.* Trans. K. Ameriks and S. Naragon. Cambridge: Cambridge University Press.

———. 2015. *Critique of Practical Reason.* Trans. M. Gregor. Cambridge: Cambridge University Press.

Katsafanas, P. 2013. *Agency and the Foundations of Ethics: Nietzschean Constitutivism.* Oxford: Oxford University Press.

Kleingeld, P. 2017. Contradiction and Kant's Formula of Universal Law. *Kant-Studien* 108(1): 89-115.

Korsgaard, Ch. 1996a. *The Sources of Normativity.* Cambridge: Cambridge University

———. 1996b. *Creating the Kingdom of Ends*. Cambridge: Cambridge university Press.

———. 2009. *Self-Constitution: Agency, Identity, and Integrity.* Oxford: Oxford University Press.

Noller, J. 2019. Reason's feeling: A systematic reconstruction of Kant's theory of moral respect. *SATS-Northern European Journal of Philosophy* 20(1): 1-18.

Nyholm, S. 2015. *Revisiting Kant's Universal Law and Humanity Formulas*. Göttingen: De Gruyter.

O'Neill, O. 1989. *Constructions of Reason.* Cambridge: Cambridge University Press.

Owen, David. 2018. Realism in Ethics and Politics: Bernard Williams, Political Theory and the Critique of Morality. In *Politics Recovered: Realist Thought in Theory and Practice*, edited by Sleat, Matt, 73-92. New York: Columbia University Press.

Reath, A. 2006. *Agency and Autonomy in Kant's Moral Theory: Selected Essays*. Oxford: Oxford University Press.

Schafer, K. 2019. Kant: Constitutivism as Capacities-first Philosophy. *Philosophical Explorations* 22(2): 177-93.

Schmitt, R. 2003. *Alienation and Freedom*. Cambridge: Westview Press.

Sensen, O. 2017. Kant's Constitutivism. In *Realism and Antirealism in Kant's Moral Philosophy: New Essays*, edited by Robinson, D.S and Elke, S., 197-222. Berlin-Boston:De Gruyter,.

Singleton, J. 2007. Kant's Account of Respect: A Bridge between Rationality and Anthropology. *Kantian Review* 12: 40-60.

Street, S. 2012. Coming to Terms with Contingency: Humean Constructivism About Practical Reason. In *Constructivism in Practical Philosophy*, edited by Lenman, J. and Shemmer, Y, 40-59. Oxford: Oxford University Press.

Tenenbaum, S. 2019. Formalism and Constitutivism in Kantian Practical Philosophy. *Philosophical Explorations* 22(2): 163-79.

Williams, B. 1995. Replies. In *World, Mind and Ethics: Essays on the ethical philosophy of Bernard Williams*, edited by J.E.. Altham, 85-224. Cambridge, Cambridge University Press.

Wood, A. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.

———. *Karl Marx*. London: Routledge.

Public Reason 14(2)-15(1): 76-82

Is Kantian Ethics Morally Alienating? Comments on Mudd

Francesco Testini
Interdisciplinary Center for Ethics, Jagiellonian University

Abstract: Kant's philosophy is notoriously based on the dichotomy between the phenomenal and the noumenal world. This dichotomy digs a rift across human nature by separating the animal and the rational parts of it, its heteronomous and autonomous components, duty and self-love. Such a dichotomy, according to Sasha Mudd, apparently gives rise to two forms of alienation: moral alienation and practical alienation. On Mudd's account, Kant successfully escapes the first kind of alienation through his doctrine of respect. Here I argue, contra Mudd, that there are at least two ways in which Kant leaves moral agents morally alienated, i.e., alienated from important dimensions of morality itself.

Keywords: objections to Kantian ethics, alienation, deontology, acting from duty.

Kant's view of human beings is, as much of his philosophy, notoriously based on the dichotomy between the phenomenal and the noumenal world. This dichotomy digs a rift across human beings by separating the animal and the rational parts of their nature, their heteronomous and autonomous components, their sense of duty and their self-love. Human beings, for Kant, inhabit both worlds.

This dichotomic view puts Kantian ethics in a sort of Cartesian predicament. Just as Cartesian metaphysics struggled to bridge the gap between *res extensa* and *res cogitans* – the thinking subject and the physical world – so Kant's moral philosophy is sometimes at pain to reconcile the tension between the two opposite dimensions of human nature. This tension manifests itself, as Sasha Mudd persuasively argued, in two specular phenomena. On the one hand, even if we agree with Kant that the fundamental moral norms are constitutive of our rational and autonomous agency and therefore inescapable, it remains unclear whether and how far these norms can motivate finite and heteronomous human beings, normally concerned with their own ends, interests and inclinations. On the other hand, although respecting moral laws may express and incarnate who we really are as rational beings, such laws may nonetheless leave us with an unfulfilled but legitimate need to reconcile morality with our own ends as finite, animal beings, inevitably concerned with living our own lives (Mudd 2023).

Mudd describes these phenomena as two forms of alienation: moral alienation (the estrangement of the heteronomous agent, motivated by happiness and inclinations, from a morality perceived as alien, foreign, and hardly motivating) and practical alienation (the estrangement of the autonomous moral agent from her empirical and heteronomous dimension – her projects, desires, aspirations, inclinations, and so on). On Mudd's account, Kant successfully escape the first kind of alienation through his doctrine of respect and attempts, unsuccessfully, to escape the second through hid doctrine of the highest good.

Here I will focus on the first phenomenon, namely moral alienation. I will suggest, *contra* Mudd, that there are at least two ways in which Kant leaves moral agents alienated from morality itself.

I. ALIENATION, SUBJECTIVE AND OBJECTIVE

Let me begin by clarifying the notion of alienation. The concept primarily applies to a problematic separation between a subject and something else (Leopold, 2018), like an object (the product of one's work, or the natural world), other persons (one's partner, family, social group), or even a component of the Self (like one's social role, fundamental projects, desires, aspirations). Notice that there is an evaluative component in this concept. The separation is problematic, and this implies that the Self and the thing it is alienated from belong together in one way or another.

There would be a lot to add, but for my purposes the following qualification will suffice. A subject can be alienated from X in two ways. It may be *objectively* alienated from X, in the sense that the relationship between the two is actually damaged, broken, or severed in some way. Or it may be *subjectively* alienated from X, in the sense that the subject experiences X as distant, detached, or separated when this is not necessarily the case (Hardimon 1994, 119-22).

These two versions of alienation are, in principle, independent from one another. Adorno, Horkheimer, and Marcuse, for instance, saw human life in an advanced capitalist society as objectively alienating, but were aware that most of the people living in these societies were feeling comfortably at home in them. On the other hand, one does not need to be objectively alienated from X to think that she is, and to suffer from this. Hegel, for instance, viewed the social and political institutions of the modern world as a place where humans should feel at home, but knew that many of his readers did not perceive them as such (Leopold, 2018)[1].

With this distinction in mind, let's now turn to moral alienation.

II. MORAL ALIENATION

Following Tenenbaum (2019), Mudd's spells out the problem of moral alienation as follows:

> The problem is that our capacity to care about and feel motivated by the constitutive end of agency is threatened by the very fact that this end seems to bypass the interests and inclinations that motivate the adoption of all our other ends. And yet, at least in principle, we need to be motivated by this end sufficiently enough for it to function as the constraint it is meant to be on our other choices. […] Alienation, on this telling, refers to the agent's failure to stand in a correct affective or motivational

1] Indeed, he partially conceived his political philosophy as a reconciliatory project, i.e., as way of showing that this was indeed the case (Hardimon, 1994).

> relation to the ends and norms to which he is, supposedly, implicitly committed just in virtue of acting at all (Mudd 2023, 65).

The problem-solving task here is an empirical and explanatory one. We are not asking how we ought to see the moral law and why we ought to obey it, but whether the moral law can motivate our actions and feature in an account for our choices. We are in the territory of moral psychology. On Mudd's account, the Kantian doctrine of respect demonstrate that we can take a pure, a priori interest in morality, thereby committing ourselves to our constitutive moral ends and thus leaving the problem of moral alienation behind us. But is it so? I believe there are at least two different ways (one objective and one subjective) in which the Kantian moral agent would be left alienated from some vital component of the moral world.

III. OBJECTIVE ALIENATION: FAR FROM OTHERS

Let's concede that the Kantian doctrine of respect can shoulder the explanatory burden by showing (I am admittedly oversimplifying this doctrine here) that human beings can actually act out of deference to the moral law and out of a sense of awe toward it. I contend that this falls short of showing that there is nothing *objectively* alienating about Kantian morality. Indeed, although the notion of respect may reconcile us with the moral and autonomous dimension of ourselves, one may argue that it does so by alienating our moral agency from the social dimension of morality itself: our immediate concern for other persons.

Kant's discourse on the foundation of morality in the *Groundwork* relegates other persons in a quite secondary role. Indeed, the appropriate object of respect, in the Kantian framework, is the moral law, not other persons. When other persons show up in Kant's discussion they usually feature as abstractions and idealizations of humanity in general, as creatures enjoying the same autonomy and subjected to the same law – not as concrete individuals, as genuinely other agents, with their own lives, plans, and claims. As Kant writes, "any respect for a person is properly only respect for the law […] of which he gives us an example" (2006[1785], 4: 402n, 14). And in his phenomenological description of respect – in the sense of awe, fear and deference the moral law instils – there is little if no trace of the care, attention, recognition and openness toward others that typically substantiate lived moral experience.

A version of such a form of alienation is at play in the difficulty that Kantians encounter in accounting for the moral weight of the agent's special relationships – loving relationships, friendships, group loyalties, and the like (a difficulty notoriously highlighted in Williams "one-thought-too-many" problem) (Williams, 1981). But the real problem runs deeper. It concerns one's moral connection with others *in general*, qua subjects with lives and experiences and concerns – not with *special* others.

Jack Samuel recently described the problem as one of social alienation in the following terms:

> Kantians […] hold that morality involves relations of mutual accountability with other people, so one might think that Kantian has therefore dodged the threat of social alienation. But these relationships of accountability can supposedly be derived from the idea of the agent as self-legislator, rather treated as *sui generis*. There is widespread skepticism that this derivation can be carried off, but even if it can, in conceiving of relations of mutual accountability as mediated by more basic self-relations, the Kantian risks social alienation. Since Kantian constructivism was in the first instance an answer to how agents recognize moral reasons, it leaves us with only indirect ways of recognizing one another. (Samuel 2022, 10)

The basic concern here is that, by constructing morality out of the constitutive features of agency and practical rationality themselves, Kantians may explain how moral reasons can motivate us, but at the price of leaving other subjects in the background (if not outside) of their image of morality, alienating moral agents from the social dimension of their moral world[2].

IV. SUBJECTIVE ALIENATION: ME, MYSELF, AND MY INTENTIONS

A second form of moral alienation is *subjective* in character. Indeed, even assuming sheer respect for the moral law can motivate human beings to act morally, and even assuming that there is nothing objectively alienating about morality, it does not follow that rational and autonomous moral agents will ever feel at home in the Kantian moral world. To the contrary, I think it is rather difficult for a reflective and sensitive moral agent to feel like she belongs to such a world.

In a nutshell, the reason is the following. One thing is to know that respect can act as a motive for a moral agent, as the Kantian doctrine of respect establishes. Another is to know that respect moved me to act in a certain case. Now, according to Kant, only an action done from the motive of duty alone, and thus out of sheer respect for the moral law, has moral worth. However, moral agents rarely (if ever) act out of duty alone, and it is difficult for them to tell when they genuinely acted out of duty. Hence it is difficult for moral agents to know whether their actions have moral worth, and thus whether they really belong, as agents, in the moral world. As a result, they may feel alienated from it.

The argument hinges on the notorious Kantian idea that only actions performed from duty have moral worth, and this idea notoriously puzzled friends and foes of Kant alike. One reason to be puzzled is the following. Situated moral agents rarely act from duty alone and seldom face the subjective limitations, obstacles, and reluctance that would allow the motive of duty to shine clearly through one's action (as it does in Kant's own examples in the *Groundwork*). This is so because inclinations, desires to live up to

2] This becomes particularly evident in the Kantians" difficulty to account for directed obligations. See Tarasenko-Struc (2020) and Samuel (forthcoming).

certain standards and to be a certain kind of persons and other heteronomous motives often team up with duty in most of the dutiful actions occurring in the world. In other words, actions conforming to morality are often overdetermined. But can these actions be morally worthy, if only actions performed from duty are morally worthy? And, when they are morally worthy, can we know that they are? Without positive answers, moral agents can hardly ever feel at home in the moral world.

Kantian scholars seem to agree that dutiful, overdetermined actions can have moral worth, but which conditions should be satisfied for these actions to have moral worth remain a debated issue, in which I cannot venture in here. Luckily, Kant himself was quite adamant about the fact that no one can ever say to know that these conditions (whatever they are, precisely) have been met on any given case. As Kant writes, "It is absolutely impossible […] to make out with complete certainty a single case in which the maxim of an action otherwise in conformity with duty rested simply on moral grounds and on the representation of one's duty" (Kant 2006, 4: 407, 19). Even after the most thorough introspection, according to Kant "one cannot show with certainty […] that the will is here determined merely through the law, without other incentive, although it seems to be so" (Kant 2006, 4: 419, 30). In other words, we might believe to have acted from duty, but we cannot claim to know that we had.

In an influential contribution, Barbara Herman tried to downplay this worry by remarking that Kant only says that we can never be certain, and that "this kind of failure no more undermines our ability to judge the motives we have acted upon than sceptical arguments undermine our ordinary judgments about ordinary objects" (Herman 1981, 370). But the analogy is misleading in at least two respects.

First, in contrast with our motives, which are accessible only through introspection, we have multiple sources of evidence about ordinary objects: at the very least, our own perceptions, others" testimony, and a scientific image of the world in which they nicely fit. Second, we have no common-sensical reasons to doubt the reliability of our perceptions about ordinary objects, whereas we have good reasons of this sort to doubt the results of our introspection into the nobility of our motives. Whereas scepticism about ordinary objects usually appeals to evil demons or brain-in-vats scenarios to put our sources of evidence into question, scepticism about the purity of our motives can rest on the documented self-indulgency and rationalizing tendencies of human psychology. In other words, both the quantity and quality of our evidence about our motives is limited, and insofar as this is the case, one may reasonably doubt of his own moral worth and feel alienated from the Kantian moral world. The kind of purity the latter requires may be simply too hard to come by.

This form of alienation is not necessarily a bad thing from a moral point of view. Indeed, one may plausibly argue that ignoring whether one's actions are truly morally worthy is desirable, as it keeps moral agents humble, reduce feeling of moral superiority and lead to a non-judgmental attitude toward fellow human beings. But it is nonetheless

a painful form of alienation, one that can make a good person with enough self-doubt feel as a moral impostor.

V. CONCLUSION

Kant's supporters may have other resources to defuse, to some extent, both these forms of moral alienation. They may draw a distinction between the moral worth of actions and that of agents or argue that moral worth is just one form of moral value (Wood 2014). And they might be right. What I hope to have showed or, at least, plausibly suggested in these few pages, is that if Kantian morality is supposed to offer "a shelter against luck" (Williams 1995, 241) – a Nietzschean point in Bernard Williams's critique of the morality system on which David Owen would probably agree[3] – then it is not a comfortable one.

francesco.testini@unimi.it

REFERENCES

Hardimon, M. O. 1994. *Hegel's Social Philosophy. The Project of Reconciliation.* Cambridge: Cambridge University Press.

Herman, B. 1981. On the value of acting from the motive of duty. *The Philosophical Review* 90(3):359-82.

Kant, I. 2006[1785]. *Groundwork of the Metaphysics of Morals.* Cambridge: Cambridge University Press.

Leopold, D. 2018. Alienation. *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/alienation/

Mudd, S. 2023. To what extent does Kant's Doctrine of the Highest Good embody a realist orientation towards ethics? *Public Reason* 14(2)-15(1): 3-18.

Owen, D. 2018. Realism in Ethics and Politics: Bernard Williams, political theory and the critique of morality. In *Politics Recovered: Realist Thought in Theory and Practice*, edited by M. Sleat, 3-92. New York: Columbia University Press.

Samuel, J. 2022. Toward a post-Kantian constructivism. *Ergo* 9(43), 1449-84.

———. (forthcoming). Alienation and the metaphysics of normativity: On the quality of our relations with the World. *Journal of Ethics and Social Philosophy*, 1-30.

Tarasenko-Struc, A. 2020. Kantian constructivism and the authority of Others. *European Journal of Philosophy* 1(28): 77-92.

Tenenbaum, S. 2019. Formalism and constitutivism in Kantian practical philosophy. *Philosophical Explorations,* 2(22): 163-79.

Williams, B. 1981. Persons, character and morality. In B. Williams, *Moral Luck*, 1-19. Cambridge: Cambridge University Press.

———. 1995. Moral Luck: A Postscript. In B. Williams, *Making Sense of Humanity*, 241-47. Cambridge: Cambridge University Press.

3] See Owen (2018).

Public Reason 14(2)-15(1): 83-98

Nietzsche's (Im)moral Psychology: Competition, Distinction, and Inequality

Matthew Bennett
University of Essex

Abstract. Honouring our influences in academia often means disagreeing with them, and with the greatest respect for David Owen's insights into Nietzsche, this is exactly what I intend to do in this paper. I will focus on one question that is prompted by Owen's Nietzsche work: how and to what degree does Nietzsche's moral psychology depart from modern moral concepts and values? The relevant point of disagreement between me and Owen is whether Nietzsche's moral psychology is compatible with modern morality. For Owen, Nietzsche expands the repertoire of moral psychology beyond the narrowly moral to the ethical. I argue instead that Nietzsche does not merely expand moral psychology. Rather, he opposes the narrowly moral with an account of self-development, freedom, and the good that is resolutely immoralist.

Keywords: Nietzsche, moral psychology, immoralism, competition, inequality.

David Owen's work has had a significant influence on my reading of Nietzsche from the very first time I tried to do any research in this area. For me, trying to understand Nietzsche's social and political thought, and in particular the possibilities therein for unorthodox ways of expanding and revising fundamental principles of liberal democracy, will always involve returning to David's work in this area. And as a result of the fact that I have been interested in Nietzsche's anti-democratic thought since my doctoral dissertation, it is no exaggeration to say that no other writer has been as consistent a presence in my engagement with Nietzsche scholarship as David.[1]

Yet honouring our influences in academia often means disagreeing with them, and with the greatest respect for David's insights into Nietzsche, this is exactly what I intend to do in this paper. David's work on Nietzsche is broad ranging and voluminous, and I cannot plausibly hope to address the virtues and limitations of all this work in one paper. Accordingly, I will focus on one question that is prompted by David's Nietzsche work: how and to what degree does Nietzsche's moral psychology depart from modern moral concepts and values? As will become clear in the sections to follow, the relevant point of disagreement between me and David is whether Nietzsche's moral psychology is compatible with modern morality. For David, Nietzsche expands the repertoire of moral psychology beyond the narrowly moral to the ethical (in a sense to be further explained in section 1). I argue instead that Nietzsche does not merely expand moral

1] The formality of referring to David by his surname seems appropriate to me in the more recognisably academic sections of this paper, but not in this introduction in which I am trying to acknowledge my debt to him. Anyone who bridles at the informality of the introduction can be assured I will return to the more familiar surname-convention in the rest of the paper.

psychology; he opposes the narrowly moral with an account of self-development, freedom, and the good that is resolutely immoralist.

I. ETHICAL, MORAL, AND IMMORAL PSYCHOLOGY

Broadly speaking, Nietzsche thought that the way we understand human thought and action is driven by what we value, or what we consider to be important, and for him that cuts across a number of different disciplines, including philosophy, history, and philology.[2] In the case of moral psychology, the values that determine its remit might be said to be built into its name, if by moral psychology we understand the study of the nature of human thought and action as it is relevant to moral evaluation. This is not to say that the conclusions of moral psychology must themselves be moral judgements, nor that they should be guided by what we might hope to be the case from the point of view of morality. But it is because morality sets the agenda for moral psychology that Nietzsche thought it would lead us to prioritise certain things at the expense of others, and to make what Nietzsche thought were mistakes about moral psychology's subject matter, such as the metaphysical nature of the will and kind of freedom needed to warrant moral blame and responsibility (GM I 13; see also Williams 1993).

By contrast, Nietzsche's alternative to moral psychology casts a wider net, encompassing psychological and behavioural insights that are not constrained by the concepts and values required for moral judgements, and that are relevant to a broader, not necessarily moral account of what it means to be a good person or to live a life worth living. In this respect we can call Nietzsche's alternative to moral psychology an *ethical* psychology. This terminological distinction follows a distinction that Owen himself has used (Owen 2002) and that he inherits from Bernard Williams. The difference in Williams's terms is this: ethics can be understood to be 'any scheme for regulating the relations between people that works through informal sanctions and internalised dispositions' (Williams 1995, 241). Williams understood morality to be one such scheme that is distinctive to a particular period in history in a particular part of the world (Williams 1985). On Williams's account there are a number of features that distinguish what he called 'the morality system' from other ethical schemes, including its obsession with the concept of obligation, its distinctive way of thinking about guilt, its refusal to accept the possibility of moral luck, and a particular way of thinking about agency that warrants accountability and blame (ibid; see also Chappell and Smyth 2023).

There is some overlap here with Nietzsche's account of what is distinctive about modern morality. On Nietzsche's analysis modern morality has a distinctive obsession with obligation, duty, and law, and it operates with a concept of guilt that is in some

2] This is one facet of Nietzsche's perspectivism, that is, his position that all claims are in some sense dependent on and expressions of idiosyncrasies of the person making the claim. Illustrative passages include D 119, GS 301, and BGE 43, 187, 231.

respects more restrictive and demanding than the kinds of reflexive negative judgements that have been important to other ethical codes (this is the subject matter of the entire second essay of GM). Modern morality also, for Nietzsche, centres some distinctive values and virtues, foremost the values of good and evil (the subject matter of the first essay of GM), and the virtues of self-denial (BGE 33), humility (HH I.137, GM I 14), and pity (A 7). And it is a slave morality, in the sense that it is an ethical scheme that uses concepts and values that develop under conditions of domination and oppression and that are developed by those who suffer the sharp end of that oppression (BGE 260).

Even this more provocative characterisation of modern morality as a slave morality can at times arguably be found in Williams's description of morality, at least insofar as Williams refers to morality as 'the peculiar institution' (Williams 1985), a euphemism used in ante-bellum United States to refer to the institution of slavery. But Williams only follows Nietzsche so far in his more hostile characterisation of morality, and there is one particularly telling point of terminology where Nietzsche and Williams do differ. Williams never goes so far as to describe himself as an immoralist; Nietzsche not only described himself this way but, as he puts it in Ecce Homo, sees his immoralism as a 'badge of honour' (EH, IV, 6).

In this respect, where Williams offers a study of ethics that insists on a broader scope than that he thought allowed by narrowly moral concerns, Nietzsche offers a study of ethics that is by design a rival to moral ethics:

> My word immoralist essentially entails two negations. First, I am negating a type of person who has been considered highest so far, the good, the benevolent, the charitable; second, I am negating a type of morality that has attained dominance and validity in the form of morality as such, - decadence morality or, to put it plainly, Christian morality. (EH, IV, 4)

That Nietzsche is hostile to Christian morality is of course no surprise to anyone who knows even a little about Nietzsche. But what is less often acknowledged is that Nietzsche's positive alternative to Christian morality is not just an ethics that is liberated from the restrictive confines of the morality system, not just a study of the virtue, freedom, and agency that comes into view when we stop obsessing over responsibility, duty, and guilt, but is rather much more radically opposed to the values of modern morality.

Nietzsche's ethical psychology is, in other words, an immoralist psychology, by which I mean that it is the study of the thought and behaviour of an ethical ideal that is incompatible with moral commitments and at times actively opposed to them. What emerges in Nietzsche's study of the psychology of this alternative, immoral higher type is a person who for anyone with moral sensibilities should seem not only strange and alien, in the way that Ajax might seem strange to a modern audience, but also objectionable, offensive, and unacceptable. And by extension, Nietzsche's own characterisation of flourishing ethical agency should be, to a modern moral reader, at

best unsettling and at worst repulsive. Or at least, this is the thesis I wish to make a case for in the rest of this paper.

I open with this rather abstract and broad-brushstroke picture of Nietzsche's ethical psychology because it is both influenced by and yet in important ways different from Owen's account of the same themes in Nietzsche. As I will recount below, albeit in an abbreviated from, Owen's work on Nietzsche rescues from him reflections on ethics and morality that tread a fine line, robustly challenging modern moral sensibilities while being acceptable to a broadly liberal readership who think that self-cultivation should be made available to all. This picture of Nietzsche is compelling, persuasive, and attractive, but I will argue that it can be so only by selecting away the more dangerous side to Nietzsche. In the sections that follow I seek to add this danger to Owen's account of Nietzsche's ethical psychology.

II. OWEN'S NIETZSCHE ON SELF-LOVE

To keep to a manageable scope I will select one theme from Owen's work on Nietzsche's ethical psychology to illustrate both its virtues and its selective attention. Owen's study of Nietzsche's ethical psychology has focused at times on what Nietzsche has to say about a collection of related reflexive, positively valanced attitudes, most notably self-love, self-respect, and self-esteem (Owen 2002, 2009). At its most general level, this work has generated a variety of important insights into the value that Nietzsche ascribed to thinking well of oneself. This work is also of particular interest for my purposes because of the way that Nietzsche's account of the nature and value of these reflexive attitudes departs from moral psychology in the narrow sense given to the term above.

Owen's account of Nietzsche on self-love is framed in part by a contrast he draws between Nietzsche and Kant. Owen observes that a central feature of Nietzsche's ethical psychology is to rehabilitate what a moralist might call egoistic attitudes, including self-love. In this respect Nietzsche differs fundamentally to Kant, at least according to Owen.[3] According to Owen, Kant denied any moral value to self-love on the same grounds that he denied moral value to the virtue of beneficence; if putatively moral actions are motivated on the basis of a generous and kind attitude either towards others or towards ourselves, such actions are vulnerable to changes in our inclinations, in what we happen to favour (Owen 2017). Owen cites a number of passages in support of this reading of Kant; I will include just one for illustration:

> This propensity to make oneself as having subjective determining grounds of choice into the objective determining grounds of the will in general can be called *self-love*; and if *self-love* makes itself lawgiving and the unconditional practical principle, it

3] I will defer to Owen on the details of Kant's thought on self-love and make no exegetical claims of my own about Kant.

> can be called *self-conceit*. Now the moral law, which alone is truly objective (namely, objective in every respect), excludes altogether the influence of self-love on the supreme practical principle and infringes without end upon self-conceit, which prescribes as laws the subjective conditions of self-love. (Kant 1999[1788], 5: 74)

Self-conceit is particularly important for the contrast Owen draws between Kant and Nietzsche, because making the 'subjective determining grounds of choice' into a law is precisely what Nietzsche at times suggests his ethical ideal would do. There are two sources that are particularly helpful for understanding how Nietzsche understands and values the activity of being a law unto oneself.

The first is Nietzsche's discussion of the sovereign individual in his genealogy of morality. Nietzsche's sovereign individual is distinguished by what he calls their 'prerogative to promise', and their capacity to deliver on self-imposed commitments that are immunised against contingencies that would undermine the steadfastness of lesser individuals (GM II 2). On Owen's account, one particularly important feature of the sovereign individual's capacity to be their own law is that the laws they follow have success conditions that are internal to acting on them, and thereby success conditions that resist formulation in advance (Owen 2009). Thus the sovereign individual is capable of committing themselves to something for the long haul, but precisely what it means for them to deliver on this commitment is only discovered in the actual delivery of the commitment.

We might wonder what promising or committing to something has to do with self-love. This is where it helps to turn to the second relevant source. In a number of key passages in *The Gay Science* (Owen focuses on *inter alia* GS 276, 334, 341, and 361; see Owen 2009) Nietzsche reflects on what it takes to not just have the ability to stick to a commitment, but to find the content of that commitment within oneself, that is, to somehow take ourselves as the source of the values and norms that govern our actions like a commitment or a promise. On Owen's account, these passages show us that part of Nietzsche's ethical ideal involves coming to love oneself through navigating the values of self-acceptance and self-creation. This involves both learning to love what is given to us by fate, an honest and loving acceptance of who we are, but also coming to love ourselves through a creative engagement with our given nature, and having the strength to create a noble and in some sense higher version of ourselves that is nonetheless realistic, without succumbing to fantasy. In this regard, Nietzsche's analysis of self-love stipulates that self-love is valuable when it has a few distinctive features: virtuous self-love both accepts what is given in our nature and also inspires us to mould that nature into something higher and noble; and virtuous self-love is put into action, it is lived, by taking the noble self we have created as a kind of ethical code, a guide for how we should live, that stands independently of the morality of the herd.

Now all of this might sit uneasily with the puritanism of Kant, or at least Owen's Kant, but it's not exactly incompatible with a more forgiving understanding of moral requirements. Perhaps coming to terms with one's own nature and treating that nature

as an ethical guide has been too often neglected in modern moral philosophy and is difficult to sustain within the requirements of the morality system, that peculiar institution. But there is nothing evidently morally impermissible in learning to love oneself, to cultivate a kind of spiritual nobility, or to commit oneself to values and goals that resist formulation and that we learn to understand only in the process of delivering them. Perhaps, then, Nietzsche's ethics includes virtues that are outside of the scope of narrowly moral thought, but not incompatible with it.

This conclusion, however, would be too hasty, for there are other features of Nietzsche's account of following one's own law, and the self-love required to do so, that are not so morally acceptable. I recount some of these features in the remaining sections below.

III. INDIVIDUALITY

The first more radically immoral feature of Nietzsche's ethical psychology comes into relief when we dig a little deeper into what Nietzsche thinks it means to follow one's own law, and in particular what makes that law genuinely one's own. For Nietzsche, the morality he opposes is among other things a herd morality, a kind of conventionalism, that Nietzsche thinks erodes individuality and demands conventional, familiar, and routine behaviour (e.g. BGE 44, 199, 202). And Nietzsche considers valuable self-love a matter of discovering and learning to love what is singular in one's own nature, and taking that singular nature as a guide to how one ought to live.

We see this sentiment expressed in some of the *Gay Science* passages cited above, specifically those in which we find Nietzsche's reflections on learning to love ourselves. In one such passage (GS 334), Nietzsche considers the way we learn to love music, and suggests we can learn lessons from doing so about love in general and love for oneself. When we develop love for music, according to Nietzsche, we must first '*learn to hear* a figure and melody at all, to detect and distinguish it, to isolate and delimit it as a life in itself' (ibid). '[T]hen', Nietzsche says, 'one needs effort and good will to *stand* it despite its strangeness; patience with its appearance and expression, and kindheartedness about its oddity'. We then, according to Nietzsche, get used to the melody, come to expect it and miss it when it's not there, until finally we become 'enraptured lovers, who no longer want anything better from the world than it and it again'.

GS 334 immediately precedes a passage entitled 'Long Live Physics', which is cited a great deal in the scholarship on Nietzsche's thought about freedom and agency (e.g. Guay 2002, Janaway 2009, Pippin 2009, May 2009). The passage is much discussed for good reason, but what is rarely noted is that Nietzsche appears to open it as a continuation of his insights about learning to love music. GS 334 is among other things about learning to patiently observe a melody in order to understand it for what it is, and learn to recognise its distinctiveness. And 'Long Live Physics' begins:

> 'How many people know how to observe! And of these few, how many observe themselves!' (GS 335)[4]

As you might imagine, Nietzsche thinks proper self-observation, the kind that allows us to learn to love what is unique to us, is very rare. But this is how he describes the rare people who are indeed capable of self-observing:

> he [who practices self-observation] would then know that there neither are nor can be actions that are all the same; that every act ever performed was done in an altogether unique and unrepeatable way, and that this will be equally true of every future act; that all prescriptions of action (even the most inward and subtle rules of all moralities so far) relate only to their rough exterior; … that our opinions about 'good' and 'noble' and 'great' can never be *proved true* by our actions because every act is unknowable; that our opinions, valuations, and tables of what is good are certainly some of the most powerful levers in the machinery of our actions, but that in each case, the law of its mechanism is unprovable. Let us therefore *limit* ourselves to the purification of our opinions and value judgements and to the *creation of tables of what is good that are new and all our own*: let us stop brooding over the 'moral values of our actions!' (GS 335, italics in original)

I think this is quite a bizarre passage, not least because there are some claims in here that at first pass seem difficult to defend.[5] But for the purposes of this paper we can restrict focus to Nietzsche's claim that successful self-observation would lead me to understand that my actions are unique and, on this basis, his recommendation that I resist the temptation to find a code of ethics that has general scope, and instead to create an ethics for myself that is unique to me, that applies to me and me alone. This is, as he puts it later in the same passage, how we 'become who we are: human beings who are new, unique, incomparable, who give themselves laws, who create themselves'.

The point, then, is that Nietzsche suggests that one of the results of learning to love ourselves is that we come to understand and appreciate our individuality, and further that the right ethical response to this individuality is to create through our actions an ethical code, a law, that applies only to us. In this respect, following one's own law, for Nietzsche, means following a law that applies only to me.

What makes this a feature of Nietzsche's immoralism? It is conceivable that one might discover and follow a law of one's own that never leads one to evidently immoral action. Perhaps the kind of person I am, or as Nietzsche might say the kind of style I give to my life, is compatible with moral values because I am by nature compassionate, dutiful, benevolent, or whatever you think a moral life must involve. But the problem with a person living by their own code, from a moralist's perspective, is that if that code is indifferent to moral obligation, then it is always possible that there will come a time

4] Translation modified. Nietzsche's German reads: 'Wie viel Menschen verstehen denn zu beobachten! Und unter den wenigen, die es verstehen, — wie viele beobachten sich selber!'

5] Consider for instance: 'our opinions about good and noble and great can never be proved true by our actions because every act is unknowable'. Why think actions are unknowable? Why think that would stop us from proving our value judgements?

when the law of one's own and the requirements of morality do come into conflict. This is also a theme in Bernard Williams's work, and also in another 20th century critic of morality, Susan Wolf, who suggests a case to illustrate this:

> Consider the case of a woman whose son has committed a crime and who must decide whether to hide him from the police. He will suffer gravely should he be caught, but unless he is caught, another innocent man will be wrongly convicted for the crime and imprisoned. I shall take it as needing no argument that impartial morality forbids protecting one's son at the expense of another innocent man's suffering. Impartial morality forbids it—but we are talking about a woman and her son. (Wolf 2014, p.39)

There are I think a number of ways of unpacking the ethical dilemma here, but there is one reading of this case that is particularly important for my purposes, which is that the dilemma here is a conflict between the requirements of morality and what it would take to continue to be myself. As Wolf herself puts it: 'if the meaning of one's life and one's very identity is bound up with someone as deeply as a mother's life is characteristically tied to her son's, why should the dictates of impartial morality be regarded as decisive?' (Wolf 2014, 41) In this sense, Wolf's case illustrates the conceivability of scenarios in which a person's identity, the things that distinguish them, that make them who they are, can generate ethical demands that do conflict with morality. And it is this possibility that makes Nietzsche's ethical ideal of following our own laws an immoralist ideal.

Nietzsche himself was of course very aware of this, and actively embraced, perhaps even at times exaggerated, the extent to which being a law unto oneself meant breaking with, and sometimes flying in the face of, moral requirements. To cite one more example, this time one of Nietzsche's own, consider the words he attributes to Napoleon, an example Nietzsche often uses of a great individual following their own law, in this case an example of world-historical proportions:

> 'I have the right to answer all charges against me with an eternal "That is me". I am apart from all the world and accept conditions from no one. I want people to submit even to my fantasies and to find it natural when I yield to this or that distraction.' That is what Napoleon once replied to his wife when she had reasons to question her husband's marital fidelity (GS 23)

This, according to Nietzsche, is how 'the seed-bearers of the future' (ibid) would see themselves and their freedom from the moral judgements of others. But I submit that, to a modern moral reader, responding in this way to the concerns of another (particularly if that other is our partner or spouse) is at the very least a less than palatable feature of Nietzsche's picture of Napoleon, and with it a feature of Nietzsche's ideal of living by one's own law that raises doubts about its compatibility with morality.

IV. SELF-ESTEEM AND INEQUALITY

A second less palatable feature of Nietzsche's account of the person who follows their own law comes to light when we ask what it takes for a person to have such a high opinion of themselves that they are confident in the legitimacy of living by a code summed up by a simple "That is me". Nietzsche often associates such high self-regard with being noble (*vornehm*; see e.g. BGE 257-296) and sometimes deliberately trades on the ambiguity of this term, referring at the same time both to those who occupy a particular class in their respective society and those who either are, or are in a position of privilege that allows them to develop, a further spiritual nobility, a kind of ethical excellence.

Figures of nobility in Nietzsche's work often see themselves as the source of what is valuable. The noble or master classes detailed in the first essay of GM (see also BGE 260) are distinguished by the fact that their ethical code takes whatever distinguishes the identity of the nobles and turns that into an ethical good, as if the nobles take so much pride in themselves that they imagine there can be no way of understanding what it is to live a good life other than to be like them. Noble types in Nietzsche's work are also those who experience what Nietzsche calls the pathos of distance (BGE 257, TI IX.37), an awareness of privilege, rank, and higher social position, that Nietzsche suggests is necessary for a person to develop the motivation to cultivate oneself into something greater. As Nietzsche puts it:

> Without *this* pathos [of distance], that *other*, more mysterious pathos could not have grown at all, that demand for new expansions of distance within the soul itself, the development of states that are increasingly high, rare, distant, tautly drawn and comprehensive, and in short, the enhancement of the type "man" (BGE 257)

On Nietzsche's account it has been the case historically that societies that have operated with clear social stratification, and moreover evident patterns of social inequality, are the kinds of societies that have allowed some people, notably those who occupy positions of privilege, to think of themselves as higher types socially in order to then further think themselves capable of not just social superiority, but ethical, perhaps even spiritual superiority. Nietzsche seems to take this to support his claim that:

> Every enhancement so far in the type "man" has been the work of an aristocratic society - and that is how it will be, again and again, since this sort of society believes in a long ladder of rank order and value distinctions between men, and in some sense needs slavery (BGE 257)

We should not be too quick to agree with Nietzsche that perfectionism requires stratified social inequality. Even if Nietzsche is right that social inequality has historically allowed some people to be proud enough to think themselves a legitimate source of an ethical code, this does not give us good enough reason to conclude that social inequality *must be* a necessary condition for such pride to emerge, still less that such inequality is a necessary condition for anyone to live by their own laws. Owen himself has pointed out that these observations do not commit Nietzsche to denying that an egalitarian society

might also find a way of democratising self-respect, and thereby democratising a form of self-love that can help us think of our own character as a legitimate source of ethical guidance (Owen 2002).

In section 5, I will push back a little against the view that Nietzsche's perfectionism does not give him reason to oppose, and might even provide reasons to support, egalitarianism. But even if we grant that Nietzsche's ethics do not commit him to favouring social inequality, he nonetheless was undeniably anti-egalitarian. Nietzsche's distaste for egalitarianism, and for related social and political movements including democracy, socialism, anarchism, and feminism, is very clear (see e.g. BGE 202), and over his later works (works after 1886), his distaste for social equality is consistent. Among Nietzsche's general expressions of disdain for egalitarianism and egalitarian social movements are at least three arguments against social equality that sit alongside his more specific claim that social superiority generates the pathos of distance needed to motivate ethical perfectionism.[6]

The first is a complaint about what Donovan Miyasaki names 'assimilation' (Miyasaki 2015), that is, the complaint that social equality results in the erasure of differences between people and encourages uniformity and conventionalism. Social movements looking to equalise and democratise are, according to Nietzsche, an expression of a herd instinct, and attempts to eliminate special privileges and rights in a social or political sense can be seen as an expression of a more fundamental hostility to distinctiveness, idiosyncrasy, and strangeness. In BGE 202, for instance, Nietzsche claims that the 'morality of herd animals' has pervaded modern Europe's political and social institutions, and in this regard 'the *democratic* movement is heir to Christianity' (BGE 202, italics in original). Similarly, Nietzsche complains in BGE 212 that the principle of equal rights leads to 'waging a joint war on everything rare, strange, privileged'. Similarly in TI IX.37 Nietzsche describes equality (*Gleichheit*) as 'a certain factual increase in similarity that the theory of 'equal rights' only gives expression to'.

Nietzsche's second case against egalitarianism is a levelling-down objection: increasing social equality inevitably involves decreasing the quality of life and character of the privileged, and in doing so erodes opportunities to generate great people and great culture. Thus Nietzsche continues his case against equal rights in BGE 212 on the grounds that the principle not only opposes 'everything rare, strange, privileged', but also 'the higher man, higher soul, higher duty, higher responsibility, […] creative power and mastery'. Egalitarianism is, for Nietzsche, a threat not just to idiosyncrasy but also to superiority of character and ability. Note that this is one feature of Nietzsche's anti-egalitarianism that is demonstrably inconsistent over his works. The same objection appears in HH 300, but it is directed not at the desire for equality *per se* but at a particular kind of desire for equality, alongside a desire to 'raise oneself and everyone else up'.

6] My categorisation of these arguments owes much to Miyasaki 2015, though our categorisations are not identical.

Sometimes, then, it seems Nietzsche does not think that egalitarianism inevitably 'levels-down'.

Third, and related to the levelling-down objection, Nietzsche sometimes claims that the cultural productivity of great individuals is possible only thanks to an unequal distributive pattern whereby the work of the many secures abundance of resources and leisure to allow a select few to be free to pursue greatness. This distributive thesis appears in *The Greek State* (I will cover this in more detail in section 5) and *inter alia* HH 439 ('A higher culture can come into existence only where there are two different castes in society: that of the workers and that of the idle, of those capable of true leisure') and BGE 258 ('the essential feature of a good, healthy aristocracy is that it does not feel that it is a function … and, consequently, that it accepts in good conscience the sacrifice of countless people who have to be pushed down and shrunk into incomplete human beings, into slaves, into tools, all for the sake of the aristocracy.')

Each of these arguments support Nietzsche's view that egalitarianism is likely to undermine the social conditions that help a privileged few work towards his perfectionist ethical ideal. Nietzsche came to believe in his later life that social equality could only be achieved at the expense of opportunities for spiritual and cultural self-development among those capable of greatness, either because it must eliminate the material conditions that allow a select few to thrive, or because egalitarian social movements tend to be hostile to individual self-development. Moreover, Nietzsche not only thought that some form of social inequality is necessary to allow some of us the pride to live by our own laws – his pathos of distance claim – but also that the erosion of this inequality would also erode any instincts some of us might have to stand apart, to break with conformity, and to create and live by an ethics of our own. In this respect, Nietzsche's ethical psychology is more closely connected to his anti-egalitarianism than we might prefer it to be.

V. COMPETITION

Nietzsche's position appears to be that only a clear social hierarchy would allow some people enough self-esteem to think themselves worthy of perfectionist self-cultivation, enough ego to think "that is me" is good enough of a reason to ignore moral and social norms. But is this really the only way that a person could think well enough of themselves to care about following their own law?

One could note that Nietzsche himself does sometimes acknowledge that herd morality can also generate a particular kind of self-confidence for a particular kind of person, specifically those who exhibit the virtues praised by herd morality, e.g. being peaceful, gentle, mild, and friendly (KSA 12.497). But this is clearly not a satisfactory option for Nietzsche's ethical psychology, for building the self-confidence of a person well-suited to herd-conformism is not going to lead to people living by their own laws.

An alternative route to generating self-esteem that has been considered by many Nietzsche readers – including Owen – is institutionalised struggle and competition, modelled on (a particular way of understanding) classical Greek agonistic culture (Acampora 2013, Connolly 1991, Hatab 1995, Owen 2002). The relevant argument in this reading begins with Nietzsche's suggestion that strength seeks to demonstrate itself (KSA 13.294), and that one of the proving grounds in which we can do this is in a struggle with opposition ('Am Meister lernen, am Gegner sich erkennen!' KSA 7.400). Competition, according to the agonistic Nietzsche, is one particularly valuable way in which we can distinguish ourselves in struggle against opposition, for competition also comes with the added benefits of channelling otherwise destructive and violent envy into a culturally generative form (a theme covered in some detail in Nietzsche's *Homer's Contest*). Thus the various forms of non-violent social competition that pervaded a period in ancient Greece[7], at least according to a position popular among some German philologists in the early 1870s, provided opportunities to develop and display physical, intellectual, rhetorical, and political prowess without having to resort to violent conflict or outright war.

The best-known focal text for Nietzsche's account of Greek agonistic culture is his unpublished essay *Homer's Contest,* but the theme also arises in the form of violent conflict in another essay that Nietzsche wrote at the same time, *The Greek State*, which paints a very different picture than that relied on by agonistic readers of Nietzsche (more on *The Greek State* in a moment). The notebooks of the period, particularly those of 1871, also contain a significant number of notes related to competition, (e.g. KSA 7.394-408) including comments on competitions between poets, sometimes specifically competition between Hesiod and Homer, and references to Sophocles as a tragedian of the suffering of agonal individuals. Moreover, and continuous with section 3's theme of distinguishing oneself as an individual, Nietzsche's notes on Greek agonistic culture from 1871 also sometimes claim that competition can perform an individuating function, helping us to step out from the crowd without overinflating our ego (e.g. KSA 7.402). And if one excels enough in competition one might find oneself not only distinguishing oneself as a cut above the crowd, but also setting the standard by which others might judge themselves. Nietzsche noted that artists in particular would strive in competition with one another to outshine all others to the point where they would set a new norm, becoming an object of imitation of the next generation (for more on this see Pearson 2022, 103).

In this regard, it seems at least the Nietzsche of 1871 would agree that institutionalised competition could function as a way of building both self-esteem and the ability to distinguish oneself as a singular source of norms and values. As we have seen above, both of these capacities can help us live up to Nietzsche's non-moral standard of the good life. Most importantly, it has seemed to some Nietzsche readers (Connolly 1991, Hatab 1995, Owen 2002) that institutionalised competition could be a mechanism that provides these

7] Or at least the account of this period in Greece popular among some prominent German philologists in the 1870s, including Nietzsche himself, Jacob Burckhardt, and Ernst Curtius.

benefits without the need for Nietzsche's immoral social and political commitments, first and foremost his inegalitarianism. Agonistic readings of Nietzsche tend to emphasise the equality between competitors needed for any institutionalised competition to function, for too much inequality in ability between competitors would mean that the weaker opponent is incapable of posing a challenge to the stronger. Along these lines it is common for agonistic readings to cite Nietzsche's claim in *Homer's Contest* that the Greeks practised ostracism in order to eliminate competitors who became too powerful (KSA 1.788), and in doing so threatened the sustainability of the relevant competition (most notably competitions for political and social status, with ostracism serving as a mechanism to prevent tyranny). This feature of a well-functioning institutionalised competition is also at least ostensibly consistent with the value that Nietzsche finds in successfully seeking competitors equal to oneself and thereby capable of posing a worthy challenge (e.g. EH I.7)

But the problem for these readings is that Nietzsche's inegalitarianism is not independent of his positive evaluation of Greek agonistic competition; the two are more closely connected than the agonistic reading would have us believe. One reason for this is indicated by Nietzsche's inclusion of both *The Greek State* and *Homer's Contest* together in the same gift-package to Cosima Wagner of 'Five prefaces to five unwritten books'. The civilised and tamed Greek competitive spirit (examined in *Homer's Contest*) co-exists with jealous violent conflict (*The Greek State)* in the same account of Greek agonistic culture that Nietzsche offers over the two essays. And Nietzsche is clear in *The Greek State* that on his account the 'genius' that Greek competition helped to develop was also a product of stark social inequality. According to *The Greek State,* the connection between the state and the production of genius in classical Greece is that the former secures peace – a familiar Hobbesian approach to state legitimacy – but also prosperity, specifically by enforcing a pyramidal social structure in which a majority works, some as slaves, to produce the material abundance required for the leisure of a minority. Such a structure ensures, Nietzsche suggests, the material conditions for a creative minority to produce culture.

And this is not for Nietzsche simply an idiosyncrasy of the Greek context, as if modern Germany might be capable of the same cultural production without the same social hierarchy and division of labour. Nietzsche directs his readers to infer a modern lesson from what he takes to be ancient wisdom:

> In order for there to be a broad, deep, fertile soil for the development of art, the overwhelming majority has to be slavishly subjected to life's necessity in the service of the minority, beyond the measure that is necessary for the individual. At their expense, through their extra work, that privileged class is to be removed from the struggle for existence, in order to produce and satisfy a new world of necessities. Accordingly, we must learn to identify as a cruel-sounding truth the fact that slavery belongs to the essence of a culture: a truth, granted, that leaves open no doubt about the absolute value of existence. This truth is the vulture which gnaws at the liver of the Promethean promoter of culture. The misery of men living a life of toil has to be increased to make the production of the world of art possible for a small number of Olympian men. (*The Greek State,* KSA 1.767)

With evidence drawn from *The Greek State* and from other passages concerning the nature and value of competition, we can identify two reasons Nietzsche gave for thinking not just that social inequality is valuable, but that (*pace* agonistic democratic readings) it is specifically a valuable feature of agonistic culture. The first is that social inequality allows a privileged class the time and leisure to engage in agonistic competition. This is a more specific application of Nietzsche's distributive anti-egalitarian thesis about the material conditions needed to allow some people – but crucially, not everyone – to flourish. It is, moreover, complementary to the analysis of the agon in *Homer's Contest*, for it is only thanks to the abundance produced by stark social inequality that a privileged class were afforded the leisure to cultivate and prove themselves in agonic contests. As James Pearson observes, Nietzsche shared this view with Jacob Burckhardt, who maintained it was only possible for Greek nobility to engage in apparently socially useless practices such as sporting contests thanks to the surplus labour of manual labourers (Pearson 2022, 92-93). Although some notes indicate that Nietzsche included labourers in his understanding of the Greeks who engaged in agon (KSA 7.396), he nonetheless considers the highest achievements of Greek culture to be the product of competitions of 'genius' that are made possible by the work of others.

The second reason Nietzsche had for thinking that social inequality is a valuable feature of agonism is that he thought competition valuable not just for its ability to incentivise self-cultivation and the pursuit of excellence, but also for its function of conferring distinction on those worthy of competition. In the Greek context, social stratification and division of labour play a role in Nietzsche's account of the distinguishing function of competition insofar as there are certain contests which only landed aristocracy are free to enter. But Nietzsche's focus on the distinguishing function of competition becomes much more prominent in later texts, particularly when he describes his own relation to his philosophical opponents. Consider for instance what Nietzsche has to say about his 'warlike nature' in EH I.7:

> One way of measuring the strength of an attacker is by looking at the sort of opponents he needs; you can always tell when something is growing because it will go looking for powerful adversaries – or problems: since a warlike philosopher will challenge problems to single combat. The task is not to conquer all obstacles in general but instead to conquer the ones where you can apply your whole strength, suppleness, and skill with weapons, - to conquer opponents that are your equals . . . Equality among enemies - first presupposition of an honest duel. You cannot wage war against things you hold in contempt; and there is no war to be waged against things you can order around, things you see as beneath you. (EH I.7)

Democratic or egalitarian readings of Nietzsche's agonism tend to emphasise the fact that Nietzsche acknowledge the need for equality between competitors; indeed, as Nietzsche puts it here, 'equality among enemies [is the] first presupposition of an honest duel'. But as this passage also illustrates, Nietzsche thought that equality among equal

competitors goes hand in hand with a discerning selection of opponents, and contempt for those unworthy of opposition (see also BGE 260 and TI IX.48).

The point for Nietzsche is, then, that competition is valuable in part because it is one means by which those who are in some sense superior can generate or reinforce their pathos of distance which, as we saw in the section 5, Nietzsche thought of as an important psychological precondition for perfectionism. A logic of inequality, Nietzsche claimed, is built into this function of competition.

VI. CONCLUSION

I do not mean to defend Nietzsche's anti-egalitarianism; I think this element of his work, indeed like quite a lot of his social and political thought more generally, is particularly crude, and much less perceptive than his analyses of religion, art, philosophy, and individual behaviour and psychology. It seems to me that particularly his claim that egalitarianism must bring with it conformity is completely off the mark. But this is nonetheless something that Nietzsche did himself think, and what I have tried to show is that his anti-egalitarianism is part of his broader immoralist account of what valuable self-love looks like, and what it takes for a person to develop the capacity to have enough self-esteem to live by an ethics that is uniquely one's own.

I have given a brief account of how Nietzsche thinks that being a law unto oneself requires living by an ethic that is at best indifferent to the requirements of morality, and is likely to at times force us to break moral rules in order to become who we are. I have also given a brief account of why Nietzsche thinks that social equality would stand in the way of anyone thinking highly enough of themselves to take a self-given ethic, a simple "This is me", to be enough justification for living by one's own law. And this in turn, for Nietzsche, justifies a blanket rejection of the value of equality, democracy, and universal rights. It is in this respect that I think that Nietzsche's account of freedom, agency, and virtue is not just an ethical but an immoral psychology.

matt.bennett@essex.ac.uk

REFERENCES

Nietzsche's work:

Cited numbers in published works refer to section numbers (e.g. GS 23 = *The Gay Science* section 23). When citing Nietzsche's unpublished work, including his notebooks, I provide the KSA reference including a volume number and page number (e.g. KSA 12.497 = KSA volume 12 p.497). KSA refers to *Sämtliche Werke: Kritische Studienausgabe in 15 Bänden* edited by Giorgio Colli and Mazzino Montinari. Berlin: de Gruyter, 1980.

Quotes in English from Nietzsche's work are taken from the following translations unless otherwise stated:

HH = *Human, All Too Human* translated by RJ Hollingdale, Cambridge University Press, 2007

WS = *The Wanderer and His Shadow* translated by RJ Hollingdale, Cambridge University Press, 2007
D = *Daybreak* translated by RJ Hollingdale, Cambridge University Press, 2006
GS = *The Gay Science* translated by Josefine Nauckhoff, Cambridge University Press, 2001.
BGE = *Beyond Good and Evil* translated by Judith Norman, Cambridge University Press, 2002
GM = *On the Genealogy of Morality* translated by Carol Diethe, Cambridge University Press, 2008
TI = *Twilight of the Idols* translated by Judith Norman, Cambridge University Press, 2007
A = *The Antichrist* translated by Judith Norman, Cambridge University Press, 2007
EH = *Ecce Homo* translated by Judith Norman, Cambridge University Press, 2007

Other works:

Acampora, Christa Davis. 2013. *Contesting Nietzsche.* Chicago: University of Chicago Press.
Chappell, Sophie-Grace and Nicholas Smyth. 2023. Bernard Williams. *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition), edited by Edward N. Zalta and Uri Nodelman. Forthcoming URL = https://plato.stanford.edu/entries/williams-bernard/.
Connolly, William. 1991. *Identity/Difference.* New York: Cornell University Press.
Guay, Robert. 2002. Nietzsche on Freedom. *European Journal of Philosophy* 10(3): 302-27.
Janaway, Christopher. 2009. Autonomy, Affect, and the Self in Nietzsche's Project of Genealogy. In *Nietzsche on Freedom and Autonomy,* edited by Ken Gemes and Simon May. Oxford University Press.
Kant, Immanuel. 1999[1788]. Critique of Practical Reason. In *The Cambridge Edition of the Works of Immanuel* Kant, translated by Mary J. *Practical Philosophy* Cambridge University Press.
May, Simon. 2009. Nihilism and the Free Self. In *Nietzsche on Freedom and Autonomy,* edited by Ken Gemes and Simon May. Oxford University Press.
Miyasaki, Donovan. 2015. The Equivocal Use of Power in Nietzsche's Failed Anti-Egalitarianism. *The Journal of Moral Philosophy* 12(1): 1-32.
Owen, David. 2002. Equality, Democracy, and Self-Respect: Reflections on Nietzsche's Agonal Perfectionism. *The Journal of Nietzsche Studies*, vol. 24, 113-31.
———. 2009.Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency. In *Nietzsche on Freedom and Autonomy,* edited by Ken Gemes and Simon May. Oxford University Press.
———. 2017. Freedom as independence: Kant and Nietzsche on non-domination, self-love and the rivalrous emotions. In *Nietzsche and Kantian Ethics*, edited by Constancio, Joao and Bailey, Tom. London. Bloomsbury Academic.
Pearson, James. 2022. *Nietzsche on Conflict, Struggle and War.* Cambridge University Press.
Pippin, Robert. 2009. How to overcome oneself: Nietzsche on freedom. In *Nietzsche on Freedom and Autonomy,* edited by Ken Gemes and Simon May. Oxford University Press.
Williams, Bernard. 1985. *Ethics and the Limits of Philosophy.* London: Fontana.
———. 1993. Nietzsche's Minimalist Moral Psychology. *European Journal of Philosophy* 1(1): 4-14.
———. 1995. Moral Luck: a postscript. In his *Making Sense of Humanity.* Cambridge: Cambridge University Press.
Wolf, Susan. 2014. Morality and Partiality. In her *The Variety of Values: Essays on Morality, Meaning, and Love* Oxford University Press.

Public Reason 14(2)-15(1): 99-104

Nietzsche's Impossible Ethics: Comments on Bennett-Owen

Previn Karian
Keele University

Abstract. This article develops themes in Matt Bennett's reservations about David Owen's endorsement of Nietzschean ethics derived from Bernard Williams. The Continental reception of Nietzsche after Heidegger is used to justify Owen's ethical hermeneutics, whilst the difficulties of Bennett's search for a moral psychology in Nietzsche's 'immoralist' are highlighted. Late Nietzsche texts are representative of a bi-conditional logic: I am strong *if and only if* you are weak. Kant's moral frame and *Anthropology* critique Nietzsche's claims to immorality, the unique law, and the sovereign individual incapable of egotistic pluralism.

Keywords: Heidegger, ethics, moral psychology, immoralist, bi-conditional logic, Kant, agonistic.

As a respondent to Matt Bennett's paper, my primary difficulty is an almost comprehensive agreement with his statements about Nietzsche. Subsequently, I merely hope to highlight some aspects that *could* (rather than *should*) have been featured in his paper.

Bennett's central concern with David Owen's writings on Nietzsche, which he otherwise highly admires, lies in the area of ethics. Bennett alleges that Owen makes use of a distinction from Bernard Williams regarding the terms "moral" and "ethical". The "moral" is viewed in characteristically derogatory terms through Nietzsche as a mindless herd or "slave" morality of compliant normativity.[1] The "ethical" is presented by Owen as a separate realm that allows for a different form of conduct, residing not in social norms but in individual choice. This relies on a set of terms around self-overcoming, self-love and "unique law" that is set free from "responsibility, duty, and guilt" located in Kantian ethics. Whilst Bennett implies a presupposition of benevolence in Owen, he finds something that he describes as "objectionable, offensive, and unacceptable" to modern moral sensibilities in Nietzsche, that lurk as a latent potential in Owen. The perceived conflict is that Owen has extracted ethical principles from Nietzsche in a manner that occludes a "danger" in the latter's "immoralist" texts. Bennett extends this conflict across themes of self-love, individuality, self-esteem, inequality and competition. He does not engage with Owen's paper published here on the "vindication-justification" distinction as an ethical base, but relies instead on a range of his other writings on Nietzsche.

It is a dominant feature of Nietzschean philosophers to make something different of his philosophy from what is found in his texts. The most striking example of this inventiveness is of course Martin Heidegger's four volume "confrontation" with Nietzsche in his lectures from 1936-40 (Heidegger 1991, 1984). Against the backdrop agenda of "rescuing Nietzsche from the Nazis", Heidegger employs his well-known

1] This is a gross caricature of both Kant and Christianity that Nietzsche necessarily misrepresents to find his own thought and voice. The misrepresentations are deferred for elaboration elsewhere.

hermeneutic method of "destructive retrieval" to find something in a philosopher which they would not have understood or seen in their work. This appears to have set a trend of "how to read Nietzsche" that is prevalent in the majority of Continental philosophers who have engaged with Nietzschean texts.[2] Heidegger's hermeneutic approach and the Continental tradition considerably mitigate Bennett's concern that Owen's Nietzsche may not be seen in the way that Bennett does. Despite his attention to textual detail,[3] Owen has no obligation to simply follow Nietzsche's thought and is at liberty to develop elements that captivate him. At its most extreme, this freedom can be found in Derrida or Bataille. Such freedom will necessarily exclude many aspects that others will find a priority in Nietzsche's texts. Bennett struggles with this freedom when applied to the ethical sphere, but this reflects his own priorities rather than Owen's.

In proposing the terms "moral psychology" and "ethical psychology", Bennett makes an adventurous move. Whilst he delineates a clear distinction that is present in Williams, his provocative contribution is to amplify the word "psychology". This shifts the ground of discussion from "values" in morals or ethics to psychological conditions in the work of value-making. The move is as perilous as it is daring, since there is a vast realm of "drive psychology" in the construction of morals and ethics that Bennett does not bring in.[4] Whilst the evident preference is to discuss Nietzsche's ethics only in terms of philosophical concepts and methodology, as Williams does, the use of the word "psychology" brings in an entirely new dimension that neither Bennett nor Owen substantively engage with. Thus, they both evade the problematic of how either "morality" or "ethics" can in any way constitute a "psychology", or whether both of those terms are purely philosophical constructs and concepts.

Bennett's further concern lies in the self-declared epithet of "immoralist" that Nietzsche ascribes to himself. He senses that "Nietzsche offers a study of ethics that is by design a rival to moral ethics." However, further clarification is required to explain why Nietzsche has any "ethics" at all rather than just another "morality". The word "immoralist" semantically refers to its opposite "moralist". Despite Nietzsche's many voluble claims of going "beyond" the moral, it is not clear that he reaches any form of construct that can be described as an "ethics", remaining stuck in an obsession with morality and its defects. This might explain the impulse in Owen and Williams to find an ethics in Nietzsche where there is none. Instead, Nietzschean texts give us the excited ravings, ecstasies and acidic nihilism of someone trapped in a hatred of morality from which he is never free, and returns to with a vengeance in his late texts. Described by Bennett as "a distinctive obsession with obligation, duty, and law…[that] operates with a concept of guilt", Nietzsche provides no alternative construct of an ethics as

2] Sample representatives of Continental philosophers who follow Heidegger's approach would be Bataille (1992), Deleuze (2006), Derrida (1978) and (1988), Klosowski (2008), and Sloterdjjk (1989).

3] See, for example, Owen 1994.

4] See Gardner 2015, Katsafanas 2016 or Meechan 2020 as a sample of current debates.

Williams does. Nietzsche's much-vaunted claim of a "revaluation of all values" results in the shattered debris of "how to philosophize with a hammer" (Nietzsche 2003c). These "reflexive negative judgments" that Bennett identifies leave us with nothing more than a negation, begging the question as to why a mere reflexivity makes something ethical. These are the ruins from which Owen extracts an optimistic ethics regarding Nietzschean self-overcoming, framed positively as self-affirmation and self-love in the content of "vindication", that finds an additional "justification" through the validation of others (Owen 1994 and 2023). The result is that Owen provides content and meaning in "worth" and "value" that reaches the edge of psychology, though not engaging with or dependent on it. In doing so, we have a much more substantive alternative to normative morality than mere Nietzschean nihilism. Ethics are visible in Owen, but deliberately and consciously unstated in Nietzsche. I defer the strong Kantian riposte to Nietzsche-Owen available in the three types of egoism classified in the *Anthropology*, particularly the conflict between "the moral egoist" and its opposite "pluralism",[5] due to space.

Likewise, it is striking that Bennett's discussion of how Owen represents "self-love" as the affirmation of a "self-care" does not engage with the compulsive need for superiority in Nietzsche.[6] The comparative conceptual setting is between Nietzsche's "self-love" and Kant's prioritisation of "duty". For the latter, reasoned maxims and the categorical imperative of an unconditioned noumenal ground[7] provide the stability and consistency of ethical possibility. This is presented as a de-prioritisation, and possible negation, of a "self-love". However, the "sovereign individual" in Nietzsche who espouses the self-appointed titles of "noble" and "free spirit" in eponymous chapters from *Beyond Good and Evil* asserts a self-valuing that is predicated on a superiority over others who must necessarily remain weak (Nietzsche 2003a). The content of this relationship is a biconditional logic whose proposition states: I am strong *if and only if* you are weak. This logic does not enter into Owen's positive optimism, and is neglected as a potential undermining of "ethics" in Bennett. However, we are left with "immorality" in Nietzsche – a conscious, calculated reaction against "herd" morality:

> The weak and ill-constituted shall perish: first principle of *our* philanthropy. And one shall help them to do so. (Nietzsche 2003c, 128)

5] Kant (2007: 240-2; § 2. 7: 128-7: 130. Owen senses the vulnerability of his ethics when he notes the conflict that can occur when vindication opposes justification; however, there is no Kantian "pluralism" available in Owen's frame which collapses in a biconditional logic discussed below.

6] This has relevance to contemporary debates on toxic masculinity in social media cult followings of figures such as Andrew Tate and his influence on male youth.

7] The Kantian unconditioned of practical reason is misrepresented in Nietzsche's *Gay Science* (335). It is a conceptualising that eludes him, originating in *CPR*'s noumena-phenomena distinction, resulting in the metaphysical demand of reason to "find the unconditioned for the conditioned cognitions of the understanding" (A307; A587/B615). Finding the "unconditioned" is the basis of "practical reason".

If this necessary determination of strong-weak in a bi-conditional logic is an ethical or psychological basis for Nietzschean value-making, Bennett needs to explain how the demand for superiority is a condition of self-love, which the positive reframing of Owen occludes, if not rejects surreptitiously. This biconditional logic remains unidentified and unexplored in Bennett's discussion in his sections entitled "Individuality", "Self-esteem and inequality" and "Competition". A specific focus can be brought to the terms "pathos of distance" and "Greek agonistic culture" that are textually not possible in Nietzsche without a strong-weak relationship and its biconditional logic.

Linked to self-love is Bennett's tentative formulation of a potential ethics based on the capacity for the "unique" that arrives through "self-observation". Referencing Nietzsche's *Gay Science*, Bennett states that such self-observation "would lead me to understand that my actions are unique and, on this basis … I resist the temptation to find a code of ethics that has general scope, and instead to create an ethics for myself that is unique to me, that applies to me and me alone." Rather than the "negative reflexivity" that he mentions above, the attempted reflexivity here is positive.

Firstly, the term "unique" in middle and late Nietzsche is an empty idealisation that merely hopes for something substantive. In these later texts, the content of hope is an *aporia* that Nietzsche's "unique" requires – the empty space each individual has to fill in for themselves. The move from void to ethics – viz, how the ethical filling in process occurs - requires explanation. Secondly, a Kantian perspective on "self-observation" from the *Anthropology* would effortlessly critique the Nietzschean notion of "self-observation", providing several possible outcomes rather than the contestable assumption of a singular "unique" (Kant 2007, 245; § 4. 7: 132-4.). Thirdly, there is no guarantee that any amount of self-observation will lead to either a unique ethics or a unique law for oneself. Psychologically, dysfunctional self-observation can lead to high states of anxiety, depression or obsessive-compulsive mental health and personality disorders. Whilst the Nietzschean position may be that dysfunctional states and their morbidity are merely the law of the weak, it is not clear that the strong escape these disorders. Once again, ethics and laws need to be clarified as either philosophical concepts or psychological products that can come from disordered or unstable minds – especially in those perceived or claiming to be strong. These three considerations problematize the terms "unique" and "self-observation" extracted from Nietzsche.

Linked to the theme of "agonistic culture" is the underdiscussed excess of verbal and ritual violence that Nietzsche licences in his "anti-egalitarian" stance. This is, of course, taken to an extreme in the work of Bataille who makes this aspect of Nietzsche brutally explicit (Bataille 2008). Bennett intuitively senses its presence when he speaks of Nietzsche's "contempt for those unworthy of opposition", and rightly challenges, if not rebuts attempts at "democratising a form of self-love" in what is an anti-egalitarian morality (not "ethics"!). It is textually clear that this feature of contempt grew into an excessively turgid stylistic feature of his prose from *Zarathustra* onwards. Nietzsche is not merely "anti-egalitarian" – he is contemptuously and malignantly so. This provides

fertile ground for research into why this occurred and what its impact would be on a purported ethics or self-love that constitutes figures such as Superman/Overman of *Zarathustra* (Nietzsche 2003 b), or the "Free Spirit" and "What is Noble?" in eponymous chapters of *Beyond Good and Evil* (Nietzsche 2003 a).

We can now see that whilst the distinction between "morality" and "ethics" might be clear in Williams and Owen, Bennett struggles to reach the latter because of his introduction of the word "psychology". Psychology has much to offer in understanding Nietzsche's fixation with "morality", but would struggle to find any "ethics" that necessarily rest on psychological conditions. Bringing psychology into ethics is fraught with epistemic danger. Both Williams and Owen discuss those terms as philosophical concepts rather than psychological modelling and terminology, which gives them an epistemic tradition and frame of reference to draw on, that Bennett does not have. This results in a lack of consistency in his use of the terms "morality" and "ethics" where it is not clear which one he is referring to in his language.

A final point to make is the danger of using the term "Nietzsche" as a global reference to all texts representing the thought of a philosopher. Given the widespread understanding of "the death of the author" proposed by Roland Barthes (Barthes 1978), it is crucial to make clear that there are dramatic semantic shifts of thought and meaning in the chronological development of a philosopher. This denies the possibility of any singularity of thought or thinking that represents "Nietzsche". Bennett's presentation of agonistic conflict as the societal conditions that produce "genius", "art" or "excellence" belong to an early Nietzsche that disappear through his middle and late works. This qualification of the cited works, *The Greek State* and *Homer's Contest*, would severely limit, if not disqualify, an attempt to claim a creative outcome, an ethics or a psychology in what evolved as the "immoralist" of late Nietzsche in *Ecce Homo*, a tragic slave to his hatred and opposition to morality.

Bennett has valiantly addressed the stalking nihilism that Owen necessarily ignores in his constructions of an ethics that frees itself from morality, initially set out by Williams. There are key points that are identified in Bennett's presentation regarding the morality-ethics distinction that he brings towards a "psychology". However, the full ravaging impact of Nietzsche's "values" that aimed to go "beyond" morality and ethics demand a more direct engagement with his psychology, and how it constructs a failed enterprise stuck intractably in the very morality it seeks to escape.

prevink@opus-psycheia.com

REFERENCES

Barthes, Roland. 1978. *Image-Music-Text*. Translated by Stephen Heath. New York: Hill and Wang.
Bataille, Georges. 2008. *On Nietzsche*. Translated by Bruce Boone. London: Continuum.
Deleuze, Gilles. 2006. *Nietzsche and Philosophy*. London: Continuum.

Derrida, Jacques. 1979. *Spurs: Nietzsche's Styles*. Translated by Barbara Harlow. Chicago: University of Chicago Press.

———. 1988. *The Ear of the Other*. Lincoln: University of Nebraska Press.

Gardner, Sebastian. 2015. Nietzsche and Freud: The 'I' and Its Drives. In *Nietzsche and the Problem of Subjectivity*, edited by Bartholomew Ryan, Maria Joao Mayer Branco, and João Constancio, 367-93. Berlin: De Gruyter.

Heidegger, Martin. 1984. *Nietzsche, Volume 2*. Translated by David Farell Krell. Harper.

———. 1991. *Nietzsche, Volume 1*. Translated by David Farell Krell. Harper.

Janaway, Christopher. 2012. Nietzsche on Morality, Drives and Human Greatness. In *Nietzsche, Naturalism, and Normativity*, 183-201. Oxford: Oxford University Press.

Kant, Immanuel. 2007. Anthropology from a Pragmatic Point of View. In *Anthropology, History, and Education*, translated by Robert Louden. Cambridge: Cambridge University Press.

Katsafanas, Paul. 2016. Value, Affect, and Drive. In *Nietzsche on Mind and Nature*, edited by Peter Kail and Manuel Dries. Oxford: Oxford University Press.

Klossowski, Pierre. 2008. *Nietzsche and the Vicious Circle*. London: Continuum.

Meechan, John. 2020. Rhythms and Drives. *Process Studies* 49(1): 79-114.

Nietzsche, Friedrich. 2003a. *Beyond Good and Evil*. Translated by Reginald John Hollingdale. London: Penguin.

———. 2003b. *Thus Spoke Zarathustra*. Translated by Reginald John Hollingdale. London: Penguin.

———. 2003c. *Twilight of the Idols* and *The Anti-Christ*. Translated by Reginald John Hollingdale. London: Penguin.

Owen, David. 2023. "On Vindication in Ethical Life" *Public Reason*.

———. 1994. *Maturity and Modernity: Nietzsche, Weber, Foucault and the Ambivalence of Reason*. London: Routledge.

Sloterdijk, Peter. 1989. *Thinker on Stage: Nietzsche's Materialism*. Translated by Jamie Owen Daniel. Minneapolis: University of Minnesota Press.

Public Reason 14(2)-15(1): 105-119

Reply to Critics

David Owen
University of Southampton

Being invited to give the Rousseau Lecture is a great honour and I am grateful to Sorin Baiasu for this invitation as well as all the tremendous work that he has done bringing together a range of critics and commentators to engage with the Kant and Nietzsche section of my rather disparate range of work that most directly related to the topic of my Rousseau lecture. In philosophy, respect often takes the form of crafting challenges to the assumptions, claims, arguments, and ends of another's work – and unsurprisingly I have learnt much from the time, care and attention that the critics and commentators have given to engaging with my writings. Because the main lines of argument and objection that they develop are quite diverse, I will address the challenges they propose one by one rather than in parallel, before trying to reflect more generally in my concluding comments.

I. SAUNDERS - THE SELF, LOVE AND THE OTHER

Joe Saunders takes up my discussion of Kant and Nietzsche on self-love, and the different strategies that each of the philosophers adopts to address the fact that self-love is a feature of our human constitution. Saunders' argument is not to directly contest the account that I offer but to target what he sees as an assumption shared by Kant and Nietzsche (perhaps rooted in their shared Protestant backgrounds) and problematically carried over into my discussion (conceived as a philosophical and not merely historical contribution), namely, that self-love is an ineliminable feature of our human make-up. For Kant, this appears as the view that the kind of finite desiring creatures that we are cannot but be concerned with our own happiness. For Nietzsche, this appears as the view that the kind of beings that we are cannot but be concerned with our power, that is, with being effective agents who can shape the world in which they find themselves. The desire for happiness for Kant and for power (effective agency) for Nietzsche are seen as important features of human psychology. Neither of these claims amounts to the view that these are the only psychological motivations that human beings have – as Saunders seems to suppose – but they do amount to the claim that these are important and abiding features of our psychology and hence need to be addressed in any plausible account of ethics. As we will see in more detail shortly, contra Saunders, neither Kant nor Nietzsche deny that "human being can be motivated by the thoughts and feelings of others" (Saunders, 21) and both take the fact that we can be so motivated as an ethically significant fact.

The core of Saunder's argument does not, however (happily), hang on the claim that Kant and Nietzsche are psychological egoists. Rather his key argument has two

central claims. The first is that we can overcome our propensity to self-love through cultivating affective empathy with others – and that we should do so. This will, Saunders argues, enable us to recognise our mutual vulnerability and dependence. The second is that while autonomy, independence, self-responsibility and self-mastery are good things, promoting these can lead us to overlook or fail to give sufficient emphasis to vulnerability, dependence and empathy – and that which of these it is important to emphasise will be indexed to social circumstances and contexts. Saunders concludes by acknowledging that Kant and Nietzsche do have moments where the values that he is concerned with come to the fore, but he presents this as "uncharacteristic" in Kant's case (24) and as underemphasised in Nietzsche's (26-28).

I have sympathy for Saunders second argument – and I think the important work of feminist philosophy has been a central corrective here. But I do have concerns about Saunders' overall argument.

The first is that, in contrast to Kant and Nietzsche, Saunders seems to be rather indiscriminate in his promotion of affective empathy which "opens us to the influence of others by causing us to resonate with their emotive states" (Aaltola cited in Sanders 23) and not perhaps alert enough to the potential dangers that empathy involves. Consider, by contrast, Kant on this issue who not only recognizes that human beings can be affected by empathy with others but argues that there are important distinctions to make in virtue of this fact:

> *Sympathetic joy* and *sadness* (*sympathia moralis*) are sensible feelings of pleasure or pain (which are therefore to be called "aesthetic" [*ästhetisch*]) at another's state of joy or sorrow (shared feeling, sympathetic feeling.) Nature has already implanted in man susceptibility to these feelings. But to use this as a means to promoting active and rational benevolence is still a particular, though only a conditional, duty. It is called the duty of *humanity* (*humanitas*) because man is regarded here not merely as a rational being but also as an animal endowed with reason. Now humanity can be located either in the *capacity* and the *will* to *share in others' feelings* (*humanitas practica*) or merely in the *susceptibility*, given by nature itself, to feel joy and sadness in common with others (*humanitas aesthetica*). The first is free, and is therefore called *sympathetic* (*communio sentiendi liberalis*); it is based on practical reason. The second is *unfree* (*communio sentiendi illiberalis,* servilis); it can be called *communicable* (since it is like the susceptibility to warmth or contagious diseases), and also compassion, since it spreads naturally among men living near one another. There is obligation only to the first. (Kant, 6: 457)

Kant's concern is that the second contagion-like mode all too easily leads to the worst kind of beneficence in which just the sort of compassion or pity at which Nietzsche also took aim is exhibited.[1] Nietzsche, for example, charges that affective empathy is liable to produce this objectionable pity because it is superficial:

1] We may note, in passing, that contemporary criticisms of humanitarianism (Fassin 2011) exhibit the same critical concern.

> That from which we suffer most profoundly and personally is almost incomprehensible and inaccessible to everyone else: in this matter we are hidden from our neighbour even when he eats at the same table with us. Everywhere, however, where we are noticed as sufferers, our suffering is interpreted in a shallow way; it belongs to the nature of the emotion of pity to divest unfamiliar suffering of its properly personal character: - our "benefactors" lower our value and volition more than our enemies. In most benefits which are conferred on the unfortunate there is something shocking in the intellectual levity with which the compassionate person plays the role of fate: he knows nothing of all the inner consequences and complications which are called misfortune for me or for you! The entire economy of my soul and its adjustment by "misfortune," the uprising of new sources and needs, the closing up of old wounds, the repudiation of whole periods of the past - none of these things which may be connected with misfortune preoccupy the dear sympathiser. He wishes to succour, and does not reflect that there is a personal necessity for misfortune; that terror, want, impoverishment, midnight watches, adventures, hazards and mistakes are as necessary to me and to you as their opposites, yea, that, to speak mystically, the path to one's own heaven always leads through the voluptuousness of one's own hell. (GS s.338)

There are further concerns that one may have that Nietzsche highlights of which I will mention one. It is that we can be all too easily overwhelmed via empathy. Nietzsche's various remarks on making oneself hard should be read not as a denial of human being's ability to be affected by the thoughts and feelings of others but precisely an acknowledgment of the fact (which, as we have seen, Kant also recognizes) that we are naturally open to being so affected. When he remarks in *The Gay Science* – and the reference here should be taken to be to, for example, military commanders or political leaders – "But not to perish from internal distress and doubt when one inflicts great suffering and hears the cry of it - that is great, that belongs to greatness" (s.325), his point is precisely that to undertake a rightful or necessary course of conduct knowing that it will cause suffering and in the face of the suffering caused requires a cultivated hardness.

Another potential line of criticism is advanced by Estrada's commentary on Saunders' essay which seeks to offer an account in which the role of empathy is not to serve as the direct motivating force of our ethical response to others. Rather empathy's role is to motivate our acknowledgment of the other as a being of dignity worthy of respect and it is then respect for the other that does the main work in terms of our direct ethical responses to the other – for example, our responses to refugees should be governed by respect for them as being entitled to dignity, not by empathy for their suffering. I am basically in agreement with Estrada on this point and I take it to be supported not only by contemporary criticisms of humanitarianism (Fassin 2011) but also by reflection on the ways in which processes of dehumanisation work in relation to groups such as refugees (Owen, 2019).

None of this is to deny Saunders' point that acknowledgment of our vulnerability and mutual dependency is important, on the contrary – and I think that his reflections

on what it is important to emphasise here and now are salient for our activity in ways I have not sufficiently appreciated.

II. SATNE - MORAL DEVELOPMENT, REPENTANCE, AND SELF-AFFIRMATION

Paula Satne's rich and careful work on Kant is manifest in her close engagement with my essay on Kant and Nietzsche on autonomy and self-love. In this essay I address the tension between love and law within the European tradition of moral philosophy as it manifests itself in Kant and Nietzsche, advancing the rather bold claim that Nietzsche has a greater claim to have resolved this tension than Kant. As Satne correctly identifies, this claim is predicated on a range of challenges that Nietzsche's philosophy poses to Kant's in this context.

Satne's previous work has drawn out an account of moral development in the Kantian agent and she summarises this effectively in order to show that there are a number of parallels between Kant's account of the morally autonomous agent and Nietzsche's account of the sovereign individual that my less developmental reading of Kant had failed to appreciate. I am happy to defer to Satne on this point. However, as she notes this simply sets up the two major challenges that Nietzsche's view (as reconstructed by me) poses to Kant's moral philosophy. The first of these concerns the distinction that I draw (along with Aaron Ridley and Robert Pippin) between Nietzsche's expressivist view of ethical autonomy in which our promises (read: ethical commitments) are such that their success conditions are, at least partly, internal to the performance of the commitment in a way that is similar to aesthetic agency. This point can be registered by distinguishing between the letter and spirit of the law (that one gives oneself) and the claim that Nietzschean autonomy requires both, whereas Kantian autonomy requires only the former. I express this point, following Ridley, by arguing that the categorical imperative in ruling out maxims that cannot be universalised serves to identify a range of unconditional "I will not" statements in which only the letter matters.

Satne argues that this is "a simplistic reading of Kant's ethics" (Satne 41) and that "each person's project of moral development is personal and shaped by features of their own psychology and personal history as well as features of their social circumstances" (Satne 8). But doesn't the point remain that the Kantian process of moral development that Satne sketches operates through revising one's maxims which are conceived as formulable independently and in advance of the actions they prescribe? Satne's argument does cause me to reconsider whether the letter/spirit objection is better seen as applying to positive commitments rather than negative prohibitions. Consider Ridley's example of marriage:

> It is true that there are some independently specifiable success-conditions here (although they are defeasible). Respect is presumably necessary, for example, as are caring for the other person's interest and not betraying them, say. But what exactly might count as betrayal, or what caring for the other person's interests might look

> like in this case – or even whether these things are what is at issue – cannot be specified independently of the particular marriage that it is, of the circumstances, history and personalities peculiar to it, and of how those things unfold or develop over time. It is, in other words, perfectly possible that everything I do is, as it were, strictly speaking respectful, considerate and loyal, and yet that I fail to be any good as a husband – I am true to the letter but miss the spirit, as we might say. (Ridley 2009, 186)

But, we might think, while the full range of what might count as betrayal cannot be specified independently of the particular marriage, this does not mean that some actions would not count as betrayal in any marriage because incompatible with the point of the institution. Ridley's focus is on the point that one can be "true to the letter but miss the spirit' but one can also fail to be true to the letter (in ways incompatible with the spirit). If this is right, then there may be scope for Satne's Kant to argue that Nietzsche's objections to the categorical imperative are illicitly conflating a procedure for testing maxims where contradiction generates unconditional "I will not"'s with a process for formulating maxims concerning, for example, what loyalty positively requires in a way that generates confusion. This creates space for a reading of Kant like Satne's to argue that Nietzsche's challenge to Kant, to the extent that it is cogent, can be met since the Kantian agent can continuously reformulate their maxim in the course of the performance of their commitment (e.g., to their marriage).

The second challenge to Kant that Satne addresses concerns Kant's framing of the distinction between morality and egoism in a way that rules out "true" self-love as the general orientation to *eudaemonia* whether conceived in Aristotelian or Nietzschean terms. Satne's strategy in addressing this challenge is, drawing on the important work of Alice Pinheiro Walla, to argue that Kant's philosophy has considerably more space for happiness than is generally recognized. I see no reason not to agree with Satne and Pinheiro Walla on this point. As they cogently argue, for Kant "it is precisely because morality is the natural end of human beings that morality and happiness are not necessarily incompatible and there can be space for the pursuit of happiness in human life', where "Kant's conception of happiness should be understood in broadly hedonist terms" (Satne, 50). But what this rules out is precisely what Aristotle and Nietzsche, in their different ways, both affirm, namely, an ethical understanding of human flourishing as the natural end of humanity in which flourishing is not conceived either in purely moral nor purely hedonist terms. Aristotle and Nietzsche both distinguish between "true" self-love and "pseudo" self-love where it is the latter that is construed in terms of a hedonist notion of happiness. Both are concerned to reject the diremption between man as a rational being and as a finite desiring creature that underlies Kant's outlook. Here then I think that while Satne's Kant can accommodate much more of a concern with happiness than many readings of this philosopher, this does not answer the challenge that Nietzsche poses. The natural strategy for the Kantian in this context though is offer a counter-challenge to Nietzsche. One reason that Kant rejects Aristotelian virtue

ethics is that such an ethical outlook hangs on a view of nature that is, Kant takes it, incompatible with modern science. Nietzsche agrees with this objection to Aristotle (and classical virtue ethics). But then Kant's reasons for treating morality as the end of humanity and for treating happiness in hedonist terms that respect the pluralism of individuals' cares and interests are perfectly intelligible, and the fundamental challenge is the one that Nietzsche confronts in attempting to persuade us of the commitment to *amor fati* as the orientation of "true" self-love (assuming that my reconstruction of Nietzsche's view is plausible). Whether Kant has any challenge to answer thus hangs on whether Nietzsche can make good his account – and that is still very much an open question.

III. MUDD - TO WHAT EXTENT DOES KANT'S DOCTRINE OF THE HIGHEST GOOD EMBODY A REALIST ORIENTATION TOWARDS ETHICS?

In her fascinating and complex essay, Sasha Mudd addresses what she will call "the problem of alienation" in Kant as a way of exploring the limits of the constitutivist reading of Kant and revealing an enduring tension between the moral standpoint construed in constitutivist terms and the good of non-alienation. Taking up an objection to constitutivist approaches to morality given acute expression by Tenenbaum (2019), she construes alienation thus:

> Alienation … refers to the agent's failure to stand in a correct affective or motivational relation to the ends and norms to which he is, supposedly, implicitly committed just in virtue of acting at all. To be alienated from the constitutive end of agency, is just, in the first instance, for the agent to fail to be motivated by this end sufficiently enough for it to function as the felt constraint it is meant to be on her choice of other ends and maxims. (65)

Her central claim is that while Kant can solve a first version of the problem of alienation (moral alienation), his philosophy remains vulnerable to a second, more subtle version (practical alienation) and indeed is made vulnerable to this second problem precisely by the framework through which he addresses the first problem. She concludes that Kant's decision to recast morality teleologically through the doctrine of the Highest Good discloses a central concern with the question of what is needed for human beings to sustain confidence in morality even if this inevitability imposes costs on us in terms of practical alienation and here, surprisingly, Kant exhibits some features of a realist outlook.

Rather than reconstructing the whole of Mudd's argument, it will suffice here to focus on two points. The first is that Mudd takes it that Kant can resolve the problem of moral alienation by showing that individuals can be motivated by respect for the moral law. The second is that Kant's way of doing this simply entrenches the distinction between our rational and animal natures, and hence the problem of practical alienation, as a constitutive part of his account of human beings:

> The point, in a nutshell, is just that insofar as the ends of happiness and morality are fundamentally heterogenous and insofar as both are equally expressive of one or the other part of our essential nature, neither moral action nor action from the principle of self-love can be considered expressive of or aligned with our true nature taken as a whole. In this sense, it seems that there is no course of action or maxim of action that can leave us fully un-alienated, for just as sure as heteronomy expresses moral alienation, so moral autonomy threatens practical alienation. (67)

One question which arises here is whether the issue of moral alienation is resolved. In his commentary Testini raises this issue:

> Let's concede that the Kantian doctrine of respect can shoulder the explanatory burden by showing (I am admittedly oversimplifying this doctrine here) that human beings can actually act out of deference to the moral law and out of a sense of awe toward it. I contend that this falls short of showing that there is nothing *objectively* alienating about Kantian morality. Indeed, although the notion of respect may reconcile us with the moral and autonomous dimension of ourselves, one may argue that it does so by alienating our moral agency from the social dimension of morality itself: our immediate concern for other persons. (78)

I think Testini has a point here, but it is one that highlights the interpersonal dimension of the issue of practical alienation and how this follows from Kant's way of resolving the problem of moral alienation. As Mudd remark's "the solution to moral alienation, action from the motive of respect, in fact creates the conditions of practical alienation, insofar as respect abstracts completely from all ends and motives associated with happiness." (68) Of course, there is a wider sense in which Testini is surely right if we think of what we want from an account of morality outside of the specifically Kantian context that Mudd is addressing.

The next step in Mudd's argument is to argue the doctrine of the Highest Good represents Kant's attempt to resolve the problem of practical alienation (rather than mitigate it which is, I take it, what the type of reading of Kant we find in Satne attempts to do). The problem is that this attempted resolution raises a critical problem for Kant's constitutivism because, given the conditions under which the Highest Good may be conceived as possible – Kant's *res fidei*, realizing the Highest Good is not something that we can rationally aim at since it is not, even in principle, something within our agential control. But then seeking the Highest Good cannot be the constitutive end of rational agency because it is not something we can intelligibly seek to realize.

Mudd's final step is to argue that Kant's reworking of morality in terms of the Highest Good might be seen as disclosing some features of a realist orientation in the way that I have, drawing on Williams, generally characterised that orientation. She acknowledges that Kant is directly at odds with realism's central methodological principle ("never explain the ethical in terms of something special to ethics if you can explain it in terms that apply to the non-ethical as well") but argues that Kant does share something of the "pessimism of strength" that characterizes the realistic outlook. Mudd makes a plausible case here in terms of Kant's rejection of the claims that the

world is fully intelligible to us, that correctly understood it will make moral sense to us, that reason necessarily contributes to happiness, or that there is fit between the exercise of reason and human flourishing. But if Kant is a pessimist in this sense, is he also a realist? Kant's doctrine of the Highest Good requires *res fidei* – free will, God and immortality of the soul – as the condition of possibility of the Highest Good. We "must" accept these postulates as matters of faith. But what is the status of this "must"? Is it a requirement of reason or of empirical psychology? Only if it is the latter would Kant's view align with the realistic outlook.

Perhaps the more important point that Mudd has developed across her whole argument, however, is one on which I think she is right in ways that matter. So Mudd points out that I situate Kant as a constitutivist and locate the more general constitutivist endeavour of which I take Kant to be an example as a response to the breakdown of the Aristotelian-type picture in which inside and outside perspectives on the value of ethical life can be reconciled through a teleological picture of nature. Mudd's argument suggests that Kant's philosophy is more complexly related to this question. I quote at length here because of the significance of this argument:

> [...] the tension between what we might think of as an inside and an outside perspective on ethics remains within Kant´s system, since it is precisely this tension that, on my reading, Kant's doctrine of the Highest Good seeks, however unsuccessfully, to resolve. That is to say, contrary to Owen, it is not the case that for Kant there is no tension between an inside and outside perspective insofar as our reasons to be moral are constitutive of or intrinsic to rational agency as such. Such a tension would plausibly disappear if we were solely pure rational beings, but instead, for Kant, we are essentially bifurcated beings, both sensible and rational. It follows that the moral norms constitutively derived from the pure rational part of our nature do not prevent us from asking justificatory questions from a different normative standpoint, namely that defined by our animal nature, with its distinctive set of ends and interests. On Kant's account, then, we are rationally compelled to step "outside" ethics in order to ask about the fate of happiness, construed as that end to which we are inescapably attached though a kind of "natural necessity". (73)

She continues:

> But, importantly, for Kant, the project of attempting to reconcile our need for happiness with the demands of morality through postulating the Highest Good as the end of pure practical reason is not undertaken for the sake of justifying morality. This is not Kant's view at all. Rather, morality and its chief law are fully justified without any appeal to ends, even that of the Highest Good. Instead, in establishing the Highest Good as the end of pure practical reason Kant is trying to re-orient us, practically, to the task that morality sets us, which is justified on independent grounds. This project of practical reorientation, in turn, displays a concern that Owen takes to characterize the realist orientation, namely a concern with "our ability to sustain confidence in the moral outlook," whatever its justificatory supports.

To put this in the slightly different terms of my Rousseau lecture, Kant is concerned to vindicate our moral outlook, to show that that we have reason to affirm our

commitment to this outlook. Of course, if Mudd is right, Kant fails to achieve this task – the problem of practical alienation is not overcome – but it is a great merit of Mudd's argument to demonstrate that the constitutivist move does not escape the problem that the collapse of the Aristotelian picture engenders. Although this is not Mudd's concern, her argument thus raises the stakes of the issue with which I concluded my response to Satne by making clear what hangs on the challenges that Kant and Nietzsche pose to one another. Nietzsche's doctrine of *amor fati* is, after all, above all else a response to the problem of practical alienation.

IV. BENNETT - NIETZSCHE'S (IM)MORAL PSYCHOLOGY: COMPETITION, DISTINCTION AND INEQUALITY

I have admired and learned from Matt Bennett's work on Nietzsche since I read his PhD thesis, and his critical reflections here provide an opportunity for me to engage with one dimension of his important work in the field. Bennett's aim is to show that the portrayal of Nietzsche as an ethical thinker that I have proposed is one that elides or ignores features of Nietzsche's ethical outlook which we have good reason to be disturbed and reject. In taking up Bennett's challenge, I hope I will be able to clarify the differences between our views.

Starting by noting, as I (and many others) have also done, some similarities across Nietzsche's and Bernard Williams' criticisms of morality , Bennett proposes that there is also an important difference between the two – "where Williams offers a study of ethics that insists on a broader scope than that he thought allowed by narrowly moral concerns, Nietzsche offers a study of ethics that is by design a rival to moral ethics" – which he locates in terms of Nietzsche's affirmation, in contrast to Williams, of the concept of "immoralism" given clear expression in *Ecce Homo*:

> My word immoralist essentially entails two negations. First, I am negating a type of person who has been considered highest so far, the good, the benevolent, the charitable; second, I am negating a type of morality that has attained dominance and validity in the form of morality as such, - decadence morality or, to put it plainly, Christian morality. (EH, IV, 4)

It is not clear to me that as much hangs on this as Bennett proposes since I take it that Williams also rejects the claim we have reason to see the good, benevolent, charitable type as the highest type of human individual and the claim to dominance of the type of (Christian) morality to which Nietzsche objects. But regardless of whether my view of Williams is right or not, the deeper question here concerns what commitment to being an immoralist entails. Consider in this context an important remark from *Daybreak*:

> It goes without saying that I do not deny - unless I am a fool - that many actions called immoral ought to be avoided and resisted, or that many called moral ought to be done and encouraged - but I think that one should be encouraged and the other avoided *for other reasons than hitherto.* (D s103)

This seems to sit uneasily with Bennett's contention that Nietzsche is committed to "an ethical ideal that is incompatible with moral commitments and at times actively opposed to them." (3) With this issue in mind, let me turn to Bennett's arguments.

After reviewing the outlines of my account of Nietzsche on autonomy and self-love, Bennett proposes to expose features elided in that account that point towards the objectionably immoralist features of Nietzsche's ethical psychology. The first that he addresses combines the point – on which we agree – that Nietzsche see the law that one gives oneself as "unique", that is, as indexed to the particular individual and as giving expression to their individuality with the observation that the law one's gives oneself and the requirements of morality can come into conflict. Let's address each and then the combination.

The idea of one's own law being unique is not a very complex thought insofar as Nietzsche is committed to the idea that some of one's commitments to values have internal success conditions. There is nothing particularly tricky about the idea that the expression of a value or norm can be individualised and indeed this seems integral to the very idea of style (a concept central to Nietzsche's work). We can recognize the individual ("signature") style of an artist, a musician, a chef, etc., and that they have reached a level where they have a distinctive style is part of their achievement. The same point applies in the context of ethical practices. It is possible to hear a thick description of an act of kindness or of courage (to take two examples) performed in one's social circle and know which of one's friends performed the act because one recognizes their personal ethical style. Nietzsche's attention to the issue of exemplars (always plural for Nietzsche) and exemplarity is based precisely on this point about giving one's own law to oneself.

It is, however, also true as Bennett argues that one's own ethical law can come into conflict with the demands of social morality. Williams' Gaughin case illustrates this point as does Wolf's example of the mother who must decide whether or not to hide her (guilty) son from the police which Bennett deploys. He remarks:

> There are I think a number of ways of unpacking the ethical dilemma here, but there is one reading of this case that is particularly important for my purposes, which is that the dilemma here is a conflict between the requirements of morality and what it would take to continue to be myself. As Wolf herself puts it: "if the meaning of one's life and one's very identity is bound up with someone as deeply as a mother's life is characteristically tied to her son's, why should the dictates of impartial morality be regarded as decisive?" (Wolf, 2014, p.41) In this sense, Wolf's case illustrates the conceivability of scenarios in which a person's identity, the things that distinguish them, that make them who they are, can generate ethical demands that do conflict with morality. And it is this possibility that makes Nietzsche's ethical ideal of following our own laws an immoralist ideal. (7)

Two points. First, if this is Bennett's argument, then it entails that, contra his earlier claim, Williams (and Wolf) is embracing immoralism in Nietzsche's sense so the distinction between them he sought to draw essentially collapses. Second, Bennett's

argument simply prejudges the question that Nietzsche and the morality critics are raising about whether so-called "impartial morality" should be seen as comprehensive and authoritative with respect to our ethical lives, whether – as it claims – "morality" just is the rational form of ethics. One point of presenting such ethical dilemmas is to say that the fact that these strike us as *dilemmas* and as *ethical* dilemmas speaks against that claim. Moreover, given that on Nietzsche's account the law that one gives oneself is not a maxim that can be stated independently and in advance of one's conduct, I have to work out what my deepest commitments are in such contexts and whether I can meaningful go on in a way that allows me to affirm living this life if I act one way or another. Such dilemmas are a (painful) source of self-knowledge concerning who and what I am, and they may be tragic in the sense that however I continue my sense of self is damaged.

Bennett may seem to be on safer ground when he turns to Nietzsche's invocation and apparent affirmation of Napoleon's grandiose claim, in response to his wife questioning his marital fidelity, that he wants to reply to any charges calling him to account with the avowal "That is me". But it matters here that we hear the silent "Take it or leave it". Napoleon's response appears, and is, arrogant and prideful, but it is not at all clear that it isn't also a way of Napoleon saying something truthful to his wife, namely, that he has sufficient self-knowledge to acknowledge that he can't honestly promise marital fidelity and if that is a red line for her then she should divorce him. To the extent that Nietzsche endorses the "That is me" type of response to being called to account by others, he is not thereby committed to the claim that this justifies the actions – e.g., marital infidelity – in question. (After all, a justified response by Europe to Napoleon's "That is me" was to exile him to St Helena.) If we consider Williams' Gaughin or Wolf' son-hiding mother and ask what is involved in them responding "That is me" to family or police seeking to call them to account, what they are saying is that they identify with and affirm their actions, that they accept responsibility for what they have done: "Here I stand, I could do no other." (as Luther put it).

Bennett's next step is to ask what are the conditions that enable the "That is me"-type response. Here he argues that Nietzsche is committed to the claim that the kind of ethical "nobility" expressed in this kind of response requires social hierarchy and that ethical nobility is available only to those at the top of the system of social stratification whose position grants them a pathos of social distance which generates a pathos of inner distance. Bennett acknowledges that I have argued that Nietzsche is not committed to the view that the cultivation of the pathos of inner distance (which is the key pathos for ethical excellence) requires the pathos of social distance and hence social hierarchy but he wants to push back against this argument – and he does so by proceeding in two steps.

The first step is to argue that Nietzsche was committed to an anti-egalitarian outlook:

> even if we grant that Nietzsche's ethics do not commit him to favouring social inequality, he nonetheless was undeniably anti-egalitarian. Nietzsche's distaste for egalitarianism, and for related social and political movements including democracy, socialism, anarchism, and feminism, is very clear (see e.g. BGE 202), and over his later works (works after 1886), his distaste for social equality is consistent. (9)

Bennett identifies three specific arguments that Nietzsche exhibits for seeing social and political egalitarianism as a threat to the kind of ethical culture that he is concerned to cultivate:

1. The "Equality => Conformity" Objection.
2. The Levelling Down Objection.
3. The Leisure Class Objection.

The first of these expresses a worry that is a common feature of nineteenth century perfectionist thinkers, so Nietzsche stands here alongside Mill, Emerson, Tocqueville, Jefferson, Thoreau, etc. Mill's call for the valuing of eccentricity and experiments in living aligns perfectly coherently with Emerson's and Nietzsche's concern with becoming what you are and giving style to your character. This concern leads into the second, a version of the levelling down objection in which the worry is that social equality means reducing the opportunities and quality of life of the privileged in a way that undermines the production of great individuals and cultural greatness. As Bennett rightly notes, Nietzsche associates this with a particular type of egalitarianism (what we might call "telic egalitarianism") according to which inequality is intrinsically bad, and sometimes opposes to it a form of egalitarianism focused on raising others up (aristocracy for everyone?). A third objection which follows from Nietzsche's opposition to telic egalitarianism relates to his view that the production of great culture requires "an unequal distributive pattern whereby the work of the many secures abundance of resources and leisure to allow a select few to be free to pursue greatness" (10).

Bennett is perfectly correct that we clearly find all of these anti-egalitarian views expressed in Nietzsche's work. The question, however, concerns whether these outlooks are necessary outcomes of his ethical immoralism. To establish this connection, Bennett needs to show that Nietzsche holds that the kind of ethical psychology he aims to promote and the production of greatness (individual or cultural) requires social hierarchy. He needs to demonstrate this because none of the preceding three arguments are sufficient to establish this conclusion. An (non-telic) egalitarian state can coherently promote individuality, realise great cultural achievements (e.g., the space programme) and public funding of the work of artists, for example.

It is worth considering, in passing, the US space programme as an example since this is undoubtedly a great cultural achievement on Nietzsche's criteria (just as the pyramids were) and while only a very few people can be astronauts, the American people can legitimately see the successful landing of a man on the moon as a collective

achievement in which they all participated (even if only through paying taxes) and which is expressive of their national values. It can thereby act as a source of meaning and value, as well as a spur to further achievement, in their individual and collective lives. There is also a moral objection to the space programme, namely, that it cost a great deal of money that could have been spent addressing poverty and suffering in the USA or elsewhere. This is the view that Nietzsche rejects.

To make his case, then, Bennett needs to show that Nietzsche's ethical psychology commits him to the necessity of social hierarchy, to the view that the pathos of inner distance requires the pathos of social distance. He tries to show this by drawing attention to the fact that the Greek agonism which Nietzsche, like Burkhardt, credits with the cultivation of individual genius and cultural greatness in Athens is, on the one hand, predicated on the equality of competitors (the relevant essay is *Homer's Contest*) and, on the other hand, on social inequality (the relevant essay is *The Greek State*). Bennett reasonably draws attention to the point that the take up of Nietzsche's agonism by contemporary "left-Nietzscheans" (such as myself) often focuses on *Homer's Contest* and ignores the anti-egalitarian argument of *The Greek State* which Nietzsche seems to reaffirm in *Beyond Good and Evil* (s.257). In addition to this point, Bennett further stresses that "Nietzsche thought that equality among equal competitors goes hand in hand with a discerning selection of opponents, and contempt for those unworthy of opposition."

There are two claims here and it is important to keep them separate. The first concerns the necessity of social hierarchy. I agree with Bennett that Nietzsche affirmed the claim that Bennett advances in early work such as *The Greek State* and may (although this is more contestable) have remained committed to it in *Beyond Good and Evil*, but I think that the connection between the pathos of social distance and that of inner distance is decisively severed by Nietzsche in the final essay of the *Genealogy of Morals* in which the ascetic priests redirection of ressentiment onto the individual develops a reflexive capacity for inner distance with no connection to social distance.[2] We become capable of orientating and disciplining our conduct by reference to an internalised ethical ordering of values expressed, for example, in the pathos of the contrast between the Overman and the Last Man. If this is right then the pathos of inner distance does not hang on social hierarchy and Nietzsche's ethics does not commit him to anti-egalitarianism even if Nietzsche, at times, mistakenly took it to do so.

What of Bennett's second point that, for Nietzsche, "equality among equal competitors goes hand in hand with a discerning selection of opponents, and contempt for those unworthy of opposition"? This is, I think, entirely right but it is not incompatible with egalitarianism. I recognize Matt Bennett as a worthy opponent who raises serious objections to my arguments and will, no doubt, have further cogent objections to my

2] I first advanced this argument in Owen 1998 and have refined and reaffirmed it in a number of later publications including Owen 2007, 2014 and 2017.

replies to his objections, whereas I do not recognize – indeed have contempt for the utterances of – "Joe Bloggs" on Twitter who is spouting nonsense about Nietzsche based on ignorance or idiocy. This does not mean that I am committed to denying the equal status of "Joe Bloggs" as a human being. Or, to give another example, having contempt for the racist or conspiracy theorist is perfectly compatible with respecting their equal standing and defending their entitlement to equal rights. A judgment about someone's ethical character and hence ethical worth in this sense does not entail a rejection of their entitlement to equal rights. Now it should be said that it is not clear to me that Nietzsche would agree with what I have just said. But my argument is not that Nietzsche did think this but that his ethical philosophy is not incompatible with this.

I am tempted by a further thought here – partly as a result of thinking about Bennett's comments – that relates to Nietzsche's agonism. The thought is this: a central problem that Nietzsche identifies with Christian morality is that it claimed sovereignty over the ethical domain; the "death of God" undercuts that claim to authority and thus makes it possible to construct an agon in which Christian morality in its secular forms is engaged in a duel with a different ethical outlook of the kind that Nietzsche reconstructs and proposes. Perhaps what Nietzsche is aiming at in his philosophical work is not the triumph of his view over its competitor if that means their utter elimination from the field of contest, but rather the construction and maintenance of the contest itself.

CONCLUSION

I hope it is at least partly clear from these replies to my critics how much I have benefitted from their engagement with my work even if I have not been able to address all their challenges or exhibit all the ways in which they have made me think again about questions I have elided, overlooked, failed to register or perceive clearly, or ways in which I have underestimated the richness of Kant or Nietzsche's work, or things I have just got wrong. It is one of the great pleasures of intellectual life that one's thinking is always open to challenge and to learning anew and differently what one thought one knew. The experience of the Rousseau lecture and the symposium that followed has been a site for such pleasure, and I am grateful to all of those involved for making it so.

dowen@soton.ac.uk

REFERENCES

Kant, I. 1996[1797]. *The Metaphysics of Morals*. Translated by M. McGregor. Cambridge University Press.

Nietzsche, F. 1997[1881]. *Daybreak*. Translated by RJ Hollingdale. Cambridge University Press.

———. 1974[1887]. *The Gay Science*. Translated by W. Kaufmann. Penguin.

Fassin. D. 2011. Humanitarian Reason. University of California Press.

Owen, D. 1998. Nietzsche, Enlightenment and the Problem of Noble Ethics. In *Nietzsche and the Future of the Human,* edited by J. Lippitt. MacMillian.

———. 2007. *Nietzsche's Genealogy of Morality.* Acumen Press (now Routledge).

———. 2014. Nietzsche's Freedom. In *Nietzsche and Political Theory,* edited by K. Ansell-Pearson. Continuum.

———. 2017. Freedom as Independence in Kant and Nietzsche. In *Nietzsche and Kantian Ethics,* edited by Tom Bailey and Joâo Constâncio. Bloomsbury.

———. 2019. Refugees and the Politics of Indignity. In *Refugees Now: Rethinking Borders, Hospitality, and Citizenship,* edited by S. Ahmed & K. Oliver, 13-26. Rowman and Littlefield.

Ridley, A. 2009. Nietzsche's intentions: what the sovereign individual promises. In *Nietzsche on freedom and autonomy,* edited by K. Gemes & S. May. Oxford University Press, 181-96.

PUBLIC REASON

Journal of Political and Moral Philosophy

Submission information: All contributions whether articles or reviews are welcome and should be sent by e-mail as attachment at publicreason@ub-filosofie.ro. Papers will be refereed on a blind basis by *Public Reason*'s referee board. Acceptance notices will be sent as soon as possible. The editors may ask for revisions of accepted manuscripts. Papers should have an abstract and key words. Please limit your paper submission to 8000 words, including footnotes and references, and format it for blind review (the text should be free of personal and institutional information). Along with the article, but not in the article, please send contact details: current affiliation, address, telephone number and email address. Authors are responsible for reinserting self identifying citations and references when manuscripts are prepared for final submission. Book reviews should have no more than 4000 words.

Manuscript preparation: All manuscripts should be formatted in Rich Text Format file (*rtf) or Microsoft Word document (*doc) with 12 pt. font size and double-spaced. Please use the author-date system documentation style, as it is presented in *The Chicago Manual of Style*. For more information about our editing rules, please go to: http://www.publicreason.ro/submission-info

Notes: The contributions have to be original and not published before. *Public Reason* does not require exclusivity in submission of articles. However, if you have submitted your article to another journal, we kindly require that you let us know about it, and notify us without any delay about other journals' decisions about your article. The work has to be approved by all co-authors. Authors wishing to use text passages, figures, tables etc. that have already been published elsewhere are required to obtain permission from the copyright owner(s). Evidence that permissions have been granted are to be sent along with the manuscripts submitted for the blind review.The authors assume all responsibility for the content published in the journal.

Book reviews: Books for review should be sent to: CSRB, Department of Philosophy, University of Bucharest, 204 Splaiul Independentei, Sect. 6, 060024, Bucharest, Romania.

Advertising Information: If you want to advertise your books, department programs, conferences, events etc. please contact us at publicreason@ub-filosofie.ro in order to get general information or information on costs, specifications, and deadlines.

Subscription and delivery: *Public Reason* is an open access e-journal. For the print version of the journal please contact us at publicreason@ub-filosofie.ro.

Permission: *Public Reason* holds the copyright of all published materials. Requests for permission to reprint material from *Public Reason* should be sent to publicreason@ub-filosofie.ro.

Online publication: *Public Reason* is available online at **http://publicreason.ro**

The journal *Public Reason* is published by KOSAK (The Keele-Oxford-St Andrews Kantian Research Centre (University of Keele)) and CSRB (The Center for the Study of Rationality and Beliefs, Faculty of Philosophy (University of Bucharest), in association with ECPR (The European Consortium for Political Research). For more information see http://www.csrc.ro/EN/home

ISSN 2065-7285

EISSN 2065-8958

ARTICLES

http://publicreason.ro
ISSN 2065-7285
EISSN 2065-8958