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.