Few concepts in the history of IR theory are as discomfiting to the contemporary mind as that of evil. Indeed, it is probably fair to say that for the better part of half a century now, successive generations of IR scholars have sought to eradicate evil (and its twin, human nature) from theoretical consideration and respectability. Even if human nature might be deemed worthy of rehabilitation by appeal to socio-biology or some other novel importation from the shiny new world modern science, evil seems beyond the pale. As a dark, reactionary, crude, and in all honesty, embarrassing inheritance from a more politically and philosophically primitive “realist” period, evil is generally seen as an idea that should be consigned to the dustiest corners of history as quickly and quietly as possible.

This paper, by contrast, seeks to take evil seriously as both a substantive position in the history of IR theory and as a potential contribution to contemporary theoretical debate – particularly the long-standing controversies surrounding realism and rationalism. In terms of the historical dimension, I suggest that the roots of modern realism’s most sophisticated engagement with evil are not to be found in the usual suspects, such as
Augustine,¹ and even less so in Machiavelli or Hobbes. On the contrary, and in contrast to his place as a quintessential liberal in IR theory, I would like to suggest that the crucial figure for thinking about evil in realism is none other than Immanuel Kant. In fact, it is Kant’s rigorous rationalism that leads to his startling claims about evil, claims which form the bedrock of what later became important strands in the classical realist theory of international politics in the twentieth century. The concern with evil in classical realism, in short, does not arise from ignorance of the power and claims of Enlightenment reason or its atavistic rejection. Nor does it reflect a lack of awareness of how Kant’s Critical philosophy demonstrated the fallacy of reducing human beings to natural determination via appeals to ‘human nature’. Instead, the realist concern with evil can be directly connected to Kantian rationalism to such a degree that, heretical as it may sound, Kant should be counted amongst the most important – if ambivalent – modern realists.

Recognizing this Kantian lineage allows us to explore with greater seriousness the place of evil in post-war realism and political theory. In the second half of the paper I seek briefly to show how this can be the case by looking at two very different figures: Reinhold Niebuhr and Hannah Arendt. Both Niebuhr and Arendt took the place of evil in politics with the utmost seriousness. Both, in their different ways, traced this concern back to Kant. And both thought that engaging with evil was essential for a cogent understanding of modern politics and the limits of modern rationalism as an account of those politics. In their eyes, ideas about evil, human nature, and about the tragic in politics, need not be reduced to crudely realist forms of naturalism or biologism. Nor are they intrinsically opposed to, or completely outside of the rationalist tradition in modern thought and social science. On the contrary, they mark a difficult shared inheritance between the two.

---

¹ As will become apparent, the post-Enlightenment view I trace certainly engages older theological and philosophical positions, but reworks them in specifically modern ways. The Augustinian view has been insightfully explored, particularly by Roger Epp; see, for example, The ‘Augustinian Moment’ in International Politics: Niebuhr, Butterfield, Wight and the Reclaiming of a Tradition, International Politics Occasional Research Paper, no. 10 (Aberystwyth, 1991). For sophisticated attempts in IR to engage substantially with evil include Renée Jeffrey, Evil and International Relations (London: Palgrave, 2007) and Renée Jeffery, ed., Confronting Evil in International Relations (London: Palgrave 2008).
humanity “is evil by nature”,
there exists a “radical innate evil in human nature”

In IR, statements such as these have a distinctly realist ring. Evil is generally connected
to theories of human nature that trace it to some kind of intrinsic animus dominandi, to
an innate will to dominate, or a violent egoism. In light of this, it is hardly surprising
modern thought has generally rebelled against the idea of evil in human nature as a form
of “essentialist” prejudice, a theological hangover from a by-gone age, or a naturalistic
fallacy - as well as condemning it as morally vacuous and even practically self-defeating.

In these arguments, contemporary rejections stand clearly and self-consciously in the
tradition of Critical philosophy inaugurated most importantly by Immanuel Kant. In the
Kantian tradition, as autonomous agents with free will, humans escape the
determinations of “nature”, whether this is conceived as prior determining causes
external to the self, or ‘internal’ causes arising from some supposedly innate, uncontrollable, and thus determining ‘human nature’. Seen in this light, the entire
purpose of the Critique of Pure Reason is in fact to demonstrate how, as a category of
reason, causality can apply to the phenomenal world of nature without applying
determinately to the world of human action which is based upon the freedom of the will.
Similarly, the second Critique, of practical reason, develops the implications of this for
freedom for agency and morality – all again based on the fundamental groundwork of
free will.

This being the case, to speak of ‘human nature’ in the sense of the natural (scientific,
causal) determination of human beings by factors external to themselves or beyond their
control is, in the most literal sense, a ‘category error’ for Kant. Similarly, to see evil as
residing in a fixed human nature would seem the anti-thesis of his entire philosophical,
moral, and political project. In this aspect of the Kantian legacy, one that is foundational
for most modern conceptions of agency, evil is revealed as an irrational or historical
prejudice, or a methodologically erroneous reification arising from confusion over the

2 For a critical appraisal, Annette Freyberg-Inan, What Moves Man: The Realist Theory of
international Relations and its Judgment of Human Nature (Albany: State University of New York
Press, 2003). In IR, efforts to engage substantially with evil include Renée Jeffrey, Evil and
International Relations (London: Palgrave, 2007) and Renée Jeffery, ed., Confronting Evil in

3 As, likewise, would be finding the source of evil in an external cause: the Devil, for
instance.
differences between the natural and human sciences. It results in a theoretically unsustainable vision of fixed or determined human subjectivity which misunderstands that the essence of humanity its capacity for agency, change, progress, and morality.

There are great strengths to these arguments, and there is no doubt that Kant’s critical philosophy provides indispensable tools for attacking essentialist views of human nature and naturalistic views of its intrinsic evil. The problem is, however, that it is only half of the Kantian inheritance, and the remainder is much more tricky. For the author of the quotations at the start of this section is none other than Kant himself. Humanity, he tells us in the most explicit terms, “is evil by nature”, and he is unflinching in his conviction that there exists “radical innate evil in human nature”. How are we to square this? Does not Kant’s insistence on the radical evil inherent in human nature contradict his claim that the essence of humanity lies in its freedom from natural determination, that individuals are responsible for their own actions, and that this provides the basis for morality?

These are difficult questions, and they have generated a range of complex responses. In this setting, I want simply to argue that the answer is no. For Kant, the existence of evil in the world is clear to anyone who is ‘realistic’ enough to see it. Like Voltaire, he has little time for those who fail to see the manifold suffering in the world and he provides his own catalogue of ills (including the machinations of states in the international system) to drive the point home. The question thus becomes, what is the source of this evil? One answer was that it was a result of ignorance. Another was that evil arose from prior circumstances that caused individuals to act in evil ways. In both of these views, evil could thus be eradicated by enlightenment in both theory and in political practice - precisely as many of Kant’s Enlightenment predecessors and contemporaries believed. Kant, however, did not. Instead, he is quite clear, in line with the argument above, that this origin cannot be found outside the will. As he puts it, “the source of evil cannot lie in

4 From Kant’s most extended treatment of radical evil in, Religion Within the Limits of Reason Alone, p.26, 27.

5 For instance, Pablo Muchnik, Kant’s Theory of Evil (Lanham, MD: Lexington Book, 2009) and Sharon Anderson-Gold and Pablo Muchnik, eds., Kant’s Anatomy of Evil (Cambridge: Cambridge University Press, 2010).

6 A position shared, of course, by many involved in field of International Relations from its founding to the present.
an object *determining* the will through inclination, nor yet in a natural impulse; it can lie only in a rule made by the will for the use of its freedom, that is, in a maxim".7

For Kant, evil is intrinsic to the will itself, to the very capacity of human beings for self-determination. To see evil as external to the individual, or as based in some kind of natural and irrepressible drives, would be to reassert the very form of naturalism that he seeks to overcome. It would be tantamount to the denial of the moral conduct that he thinks is accessible to the reason of all individuals, which cannot be ascribed to any factors external to the individual. To do so, would be to reduce the individual to natural determination, to his or her actions being caused by something external to themselves. This would not only render morality and freedom impossible; it would also engage us in an infinite regression as we searched for the causes of evil which always, logically, could have a prior cause. Instead, Kant defines evil as the conscious choice to act contrary to the moral law and the categorical imperative. For evil to exist in a fully human sense, therefore, it must – for Kant – be a product of the will itself: that is, the capacity for evil must be part of the very nature of human beings as autonomous agents with free will. Evil *as human nature is* freedom – the freedom to deny the categorical imperative, which arises from the same source in the freedom of the will. As one study has put it: “the theory of radical evil suggests that what also makes us human is our ability to oppose not only the ends imposed on us by nature, but also the ends that our own reason would freely give to itself".8

Kant identifies three forms of evil in human nature. The first, “fraility” refers to a the weakness in observing the maxims adopted on the basis of the Categorical Imperative, while the second, which he terms “impurity”, involves “the propensity for mixing unmoral with moral motivating causes (even when it is done with good intent and under maxims of the good)".9 These two forms are relatively easy to accommodate within conventional traditions of IR – indeed they provide much of the logic for the debates between liberals and their realist opponents throughout much of the history of the field – albeit with Kant often reduced to a one-sidedly naïve proponent of the first in ways that either overlook his concern with evil or seek to overcome it by equally naïve appeals to his supposed

---

7 Immanuel Kant, *Religion Within the Limits of Reason Alone*, p.17.


claim that history guarantees in some naturalistic way that good will arise from evil as conflict gradually teaches rational individuals to accept the logic of perpetual peace.\textsuperscript{10}

But it is the third possibility, the propensity to adopt evil maxims, that is, the “wickedness” of human nature or of the human heart” which is most challenging in Kant’s account of radical evil, and which stands most fully outside both rationalist and critical theories of IR. For Kant, this wickedness

or, if you like, the corruption of the human heart is the propensity of the will to maxims which neglect the incentives springing from the moral law in favour of others which are not moral. It may also be called the perversity of the human heart, for it reverses the ethical order (of priority) among the incentives of a free will; and although conduct which is lawfully good (i.e. legal) may be found with it, yet the cast of mind is thereby corrupted at its root (so far as the moral disposition is concerned), and the man is hence designated as evil.

It will be remarked that this propensity to evil is here ascribed (as regards conduct) to men in general, even to the best of them; this must be the case if it is to be proven that the propensity to evil in mankind is universal, or, what here comes to the same thing, that is it woven into human nature.\textsuperscript{11}

For Kant, the judgement of wickedness holds even when individuals do what they believe or feel to be right, if it conflicts with – or derives from sources other than - the moral dictates of reason. We can never know with rational certainty that in acting we are in fact acting morally. Moreover, wickedness cannot be reduced simply to a variation on “frailty” or even “impurity”. Although he is often accused of associating evil narrowly with sensuous or bodily motivations or desires, doing so misses what is truly radical in Kant’s vision of radical evil, which is that it is ultimately inscrutable: “No one can say with certainty”, he avers, “why this man becomes good, that man evil”.\textsuperscript{12}

\textsuperscript{10} These readings inevitably overlook Kant’s clear statement that we may have “faith” that this may be the case, and use it to work towards such an end, but that we cannot know it. There are parallels here with Booth’s focus on “hope” as playing such a role, though his secularized view of it may well stand in important tension with Kant’s view, since it can be argued that the latter’s position only makes sense within a wider belief (faith) in which the ends of God and humanity are in some way aligned. For such an argument about the Kant’s need for ultimate redemption in the face of evil, see Michalson, \textit{Fallen Freedom}, part two. For an argument that Kant’s inability to do so is one of the mainsprings of Hegel’s attempt to do so by seeing the evils of history as the working out of a progressive Spirit whereby in the words of the \textit{Phenomenology} “The wounds of Spirit heal, and leave no scars behind”, see Bernstein, \textit{Radical Evil}, pp. 46-75. For the most important, though largely contrasting, account in IR, see Andrew Linklater, \textit{Men and Citizens in the Theory of International Relations}.

\textsuperscript{11} Kant, quoted in Bernstein, \textit{Radical Evil}, p.29.

\textsuperscript{12} Kant, quoted in Michalson, \textit{Fallen Freedom}, p.30.
To restate the case, for Kant evil is a constant possibility, a categorical inescapability: the possibility of the good is the condition of the possibility of evil. But whereas the good can be rationally explicated, evil lies in the realm of inscrutability: at the limit of reason. In the forms of classical realism that follow from this logic, the problematic of realist politics is not naturalism in the sense of scientific determinism and predictability, nor is it just “self-interest” perverting moral action – though this is clearly a major theme, albeit much more complexly rendered than is often acknowledged. The challenge is the unpredictability of human action: the capacity of evil to disrupt agreed norms and patterns for reasons that remain inscrutable and thus unpredictable – including in many cases for the agent that undertakes them. This is the most radical dimension of realism – embedded in a view of evil and human nature that does not deny freedom, or rationality, or morality, but that is based upon taking them to their most radical limits. Kant, seen in this light, is not just the brilliantly systematic founder of a powerful rationalist liberalism, he is - at the very same time – one of (perhaps even the) foundational thinkers of modern Realism.

These issues and lineages may well make little sense within the conceptual distinctions and paradigmatic divisions of IR theory today. But they, and their concern with a sophisticated understanding of evil, were at the heart of classical realism. As Reinhold Neibuhr, the thinker who was perhaps unsurprisingly most attuned to these questions once put it in one of his most succinct formulations: “The effect of human freedom upon man’s social impulse and existence is the source of contractory theories of human behaviour. Realists emphasize the disruptive effect of human freedom on the community”

For Niebuhr, modern rationalism has have led to either the reduction of human freedom to natural determination via naïve conceptions of (social) science, to an over-estimation of human beings to order the world in line with a rational model, or to their opposite – a vitalist, Romantic reaction against rationalism in the name of a return to a true human

---

13 As Bernstein argues, Kant here initiates a tradition of thinking about evil that later includes Freud and, as discussed below, Arendt.

‘nature’ which can only lead to conflict and destruction.\(^{15}\) Equally, the desire of human beings to find a determinate cause (and thus a potential solution) to evil leads them to misjudge the depth of the challenge, a point worth quoting at some length:

The inclination of modern man to find the source of evil in this live in some particular event in history or some specific historical corruption is a natural consequence of his view of himself in a simple one-dimensional history. But this modern error merely accentuate a perennial tendency of the human heart to attribute wrong-doing to temptation and thus to escape responsibility for it...The important point in all these explanations is that they fail to explain why the particular sources of evil in history, bad priests, evil rulers, and ruling classes, should have had the power and the inclination to introduce evil into history... Each...has the virtue of throwing light upon the character of particular social evils and may point the way to their mitigation or elimination. But none of them explain how an evil which does not exist in nature could have arisen in human history.\(^{16}\)

Significantly, Neibuhr’s engagement with Kant only reinforces the radical nature of evil in realist theory. On the one hand, he criticises Kant for dividing mind and body, reason and impulse, and for seeming to identify evil with the latter.\(^{17}\) This, he argues, simply reproduces a destructive dualism in Western thought between a transcendent rationalism which devalues worldly motives and experience, and a simple-minded naturalism that denies the power of reason and free agency.\(^{18}\) In this, he holds, Kant “exhibits the moral complacency of the rational man”. Yet at the height of this critique, Niebuhr revealingly notes that the exception to this complacency lies in the Kant’s theory of “radical evil”, where “In analysing the human capacity for self-deception and its ability to make the worse appear the better reason for the sake of providing a moral façade for selfish action, Kant penetrates into spiritual intricacies and mysteries to which he seems to remain completely blind in his *Critique of Pure Reason*.\(^{19}\) For Niebuhr, Kant’s recognition of radical evil is essential, even if it comes at the cost of “completely shattering” the Kantian philosophical system since it marks the limits of rationalism itself, and demonstrates the limits of its ability to provide guidance in human affairs. Rationalists, he argues, identify the world with reason and the good (even if only

\(^{15}\) Both, he thinks, lead to the valorization of the state, and thus to an intensification of international conflict, as witnessed in the twentieth century.


\(^{17}\) Strikingly, he denies that this view, so often associated with Christianity, has any substantial purchase in the biblical tradition.


\(^{19}\) Niebuhr, *The Nature and Destiny of Man*, p.120, footnote 12.
ultimately via history). Kantian radical evil excepted, they cannot see freedom’s relation to evil: “they are never able to define sin as spiritual because they regard spirit as essentially good. They fail to see the paradox of evil arising out of freedom not as an essential or necessary consequence but as an alogical fact”.20 Here lies perhaps the most difficult and yet foundational claim arising from the sophisticated vision of evil as “radical evil” in classical realism: it arises from a deep understanding of the claims and the limits of both naturalism (empiricism/positivism) and rationalism, and yet then self-consciously declares that the evil arising from freedom is, in Niebuhr’s striking phrase, “an alogical fact”.21 Given the breadth of this challenge, it is perhaps not difficult to understand why it has little place in a world of IR theory which has few tools for its comprehension.

Arendt
The second thinker I want to explore is Hannah Arendt. Arendt is rarely viewed as a realist, and nor should she be – although her close intellectual on (and personal connections) to Hans Morgenthau are well-known. But Arendt is important for demonstrating the centrality occupied by a complex understanding of evil in post-war political thinking. Moreover, if Niebuhr’s often theological language adds even further grit to the grinding of teeth that generally accompanies any attempt to discuss evil in IR today, Arendt provides an apparently more secular and thus ammenable approach. Even if, as I will argue, this is deceptive, Arendt too should stand as a warning against the error of casting a concern with evil summarily into some kind of anti-deluvian pit, as is so often the temptation today.

20 Niebuhr, The Nature and Destiny of Man, p.120.

21 In relation to debates in IR theory today, the position of this form of realism might best be expressed in this rather lengthy passage: “The modern mind interprets man as either essentially reason, without being able to do justice to his non-rational vitalities, or as essentially vitality without appreciating the extent of his rational freedom…Its inability to estimate the evil in man realistically is partly due to the failure of modern culture to see man in his full stature of self-transcendence. The naturalist sees human freedom as little more than the freedom of homo faber and fails to appreciate to what degree the human spirit breaks and remakes the harmonies of unities of nature. The idealist, identifying freedom with reason and failing to appreciate that freedom rises above reason, imagines that the freedom of man is secure, in the mind’s impetus toward coherence and synthesis. Neither naturalism nor idealism can understand that man is free enough to violate both the necessities of nature and the logical systems of reason”. Niebuhr, The Nature and Destiny of Man, pp.123-24.
For Arendt, evil was central to understanding modern politics. As she famously wrote to Karl Jaspers in the wake of World War Two and the Holocaust, and with conscious reference to Kant: “evil has proven more radical than expected”. And she was clear in her conviction that “The problem of evil will be the fundamental question of postwar intellectual life in Europe” ( ). Indeed, Arendt not only referred directly to Kant’s conception of radical evil, she appears to have understood far better than most that its radical nature lay precisely in its inscrutability and in its resistance to being reduced to simplistic models of selfishness or self-regard.

Arendt is sometimes seen as differing from Kant in her controversial declaration of the banality of evil. As she famously characterized Eichmann, evil in this case “was not Iago and not Macbeth, and nothing would have been farther from his mind than to determine with Richard III ‘to prove a villain’. Except for an extraordinary diligence in looking out for his personal advancement, he had no motives at all...He merely, to put the matter colloquially, never realized what he was doing”. Although Arendt herself is at some points unhelpfully unclear on the relationship between “radical” and “banal” evil, it seems to me that she understood far better than most that the truly radical nature of Kant’s conception of radical evil lay precisely in its inscrutability and in its resistance to being reduced to simplistic models of selfishness or self-regard. As she put it in the introduction to *The Origins of Totalitarianism*, “And if it is true that in the final stages of totalitarianism an absolute evil appears (absolute because it can no longer be deduced from humanly comprehensible motives), it is also true that without it we might never have known the truly radical nature of Evil”. While evil may verge on the banal through routinization or its infliction through increasingly distant and de-personalized structures of harm, the claim that, like Kant’s radical evil, it is irreducible to selfish motives in the


23 Quoted in Bernstein, *Radical Evil*, p.206; see also his useful summary of five key points of her thinking about evil on p.208.


conventional sense remains one of the most troubling aspects of Arendt’s thinking, and that which connects it most directly and uncomfortably to political realism. Evil here appears as the *disregard* of the good, rather than its active, ‘demonic’ denial. Its inscrutability lay precisely in its lack of obviously ‘bad’ motives or self-interested behaviour. And for that very reason, radical evil was an essential concept of understanding, not simply a form of political rhetoric or a crude relic of the pre-modern age. This is the possibility that haunts Arendt, and whose social and political dynamics she was at pains to trace.

**Conclusion**

My goal in this paper has not been to defend these ideas of radical evil, but to show how evil needs to be redeemed as a substantial category and historical inheritance in IR theory, not as something to be ashamed of. For better or worse, the realist vision of evil and human nature is a consequence of its engagement with the very ideas – indeed the very theoretical traditions which it is consistently seen as opposed to: in particular, Kantian rationalism and moral theory. Realist theories that focus on the purported evil of human nature have represented exactly the kinds of naturalism, essentialism, and fatalism that IR theory wants to leave behind, and with them the forms of realism that rely on so crudely biological or pre-Enlightenment theological foundations. And so, in most cases it should. But evil, I hope to have suggested, is historically much more interesting than this allows, and an enquiry into it can enrich our understanding of the history of IR theory and of contemporary debates between realism and its competitors.

Perhaps paradoxically, the key figure here is Kant. Kant seeks to ground morality in the freedom of the will. It is the capacity of human beings for self-determination – to be free from determination by any prior cause or given factor – that is both the definition of what it is to be human and the ground for humanity’s capacity for moral knowledge and practical action based upon that moral knowledge. This, of course, is that Kant of most moral philosophy, and the Kant that has been claimed by both rationalist and critical theories of IR. Yet, unfortunately, for Kant (and perhaps for us), the issue cannot quite be left at this point because the freedom of the will is also for him the basis of the most fundamental form of evil, that which he famously terms “radical evil”.

---

27 On the latter, see particularly Ken Booth, *Theory of World Security*, pp.82-88.
Yet there is a different way of seeing the Kantian legacy, as one of inexorable tension and tragedy, which cannot be solved by the critique of scientism and positivism and the logic of the practical reason and the categorical imperative. As Michalson has argued, it may be that the key issue is not between freedom and necessity as natural determinism; instead “a deeper and more troubling tension in Kant’s world is that between freedom and something more like original sin than Newtonian necessity. Freedom is more fundamentally in conflict with itself than it is threatened by a mechanistic universe”. And this tension “is more existentially urgent and intrinsically tragic, for it is produced by freedom itself as a problem for itself”, it “is something we create…”.

Once radical evil is introduced into the equation, the weight that Kant places on humanity’s asocial sociability in *Perpetual Peace* is not enough to guarantee progress. Rationally viewed, it may have the potential to help the process, and it may even be an element that helps us to have faith that history may be moving in this direction and so to act to help realize it – a position more sensitive to Kant’s views than rationalist appropriations of Kant (and their critics) have grasped. But that faith is not only beyond rational knowledge in the same way as the knowledge of God is in the Kantian universe: it must contend with the potential for radical evil that resides in precisely the same source – the freedom of the will. Seen from this perspective, Kant’s faith runs seriously aground. There can be no guarantee from either nature or reason that rational freedom will triumph. Instead, as Michalson has compellingly argued, Kant’s rationality ultimately must rest on a belief that there exists some form of correspondence between human potential and a divine design – however impossible it may be to derive that faith from purely rational premises and deductions.

---