

Cambridge University Press
 978-1-107-41473-0 - Kant's Anatomy of Evil
 Edited by Sharon Anderson-Gold and Pablo Muchnik
 Excerpt
[More information](#)

Introduction

Sharon Anderson-Gold and Pablo Muchnik

Contemporary debates in moral philosophy have primarily been focused on meta-ethical questions about the justification of morality, disregarding the ease with which perfectly justified norms are displaced by non-moral considerations.¹ Given the scope, magnitude, and inventiveness of human wrongdoing, this philosophical trend seems utterly misguided. The challenge does not lie so much in how to justify morality, but in understanding how perfectly justified judgments are so easily disregarded by self-serving calculations.²

Kant's doctrine of radical evil has much to tell us about this. Against the widespread tendency to explain evil in terms of the pernicious power of natural inclinations, Kant believed that evil represented "an invisible enemy, one who hides behind reason and hence [is] all the more dangerous" (R 6: 57). The enemy is invisible, for "no matter how far back we direct our attention to our moral state, we find that this state is no longer *res integra*" (R 6: 58n.). And it is exceptionally dangerous, for the corruption in question is self-imposed: "genuine evil consists in our *will* not to resist the inclinations when they invite transgression" (ibid.). Since this type of volition rests on a maxim, and maxim formation in Kant always takes place under the constraints of

¹ See Otfried Höffe, "Ein Thema wiedergewinnen: Kant über das Böse," in O. Höffe und A. Pieper (eds.), *Über das Wesen der menschlichen Freiheit* (Berlin: Akademie Verlag, 1995), pp. 11–34.

² See Pablo Muchnik, "Kant on the Sources of Evil," in *Proceedings of the Tenth International Kant Congress* (Berlin: Walter de Gruyter, 2009), pp. 287–97.

the categorical imperative, evil hides at the heart of practical reason: it is the deliberate attempt to subordinate what we ought to do in favor of what pleases us. This subordination entails a reversal of the moral order of priority between the incentives in the human will: “self-love and their inclinations [become] the condition of compliance with the moral law – whereas it is the latter that, as the *supreme condition* of the satisfaction of the former, should have been incorporated into the universal maxim of the power of choice as the sole incentive” (R 6: 36).

As a result of the excessive influence of the *Groundwork* in the Anglo-American reception of Kant, however, Kant’s reflections on evil have been largely ignored in the secondary literature. Kant’s optimistic thesis about the analyticity of freedom and morality, by which the autonomous will (*Wille*) is equated to practical reason, has been mistakenly taken as Kant’s last word regarding human freedom.³ This view overlooks Kant’s gloomier reflections about the inextirpable propensity to evil in human nature, for which we are nonetheless responsible.

This collection of essays is an effort to set the record straight. Its primary goal is to explore the intellectual resources available in Kant for dealing with the question of evil. It places Kant’s views in the context of the critical system, interprets some of Kant’s most controversial assumptions, and extends his conception in novel ways to deal with urgent contemporary issues. There is more at stake, however, than settling a family dispute among Kantians here: acknowledging the promptness with which human beings are willing to neglect the claims of morality invites an account of human motivation and agency in which a robust conception of evil plays a central role. This is an invitation contemporary moral philosophers should not refuse. By making Kant’s conception of evil more available, we hope to contribute (if only indirectly) to an overdue shift in philosophical attention.

I

The anthology opens with Philip Rossi’s essay, “Kant’s ‘Metaphysics of Permanent Rupture’: Radical Evil and the Unity of Reason.” Following

³ For the seeds of this common misunderstanding, see, e.g., M. Kosch, *Freedom and Reason in Kant, Schelling, and Kierkegaard* (Oxford: Clarendon Press, 2006), H.E. Allison, *Kant’s Theory of Freedom* (Cambridge University Press, 1990), and G. Prauss, *Kant über Freiheit als Autonomie* (Frankfurt am Main: Vittorio Klostermann, 1983).

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)*Introduction*

3

Susan Neiman, Rossi argues that Kant's philosophy is not merely a response to certain epistemological and metaphysical questions (i.e., how are a priori synthetic judgments possible).⁴ More importantly, it is a response to the presence of evil, which threatens the very intelligibility of the world and our need to feel at home in it. Evil makes manifest a rift between the world as *it is* and the world as it *ought to be*, inciting us to find unity and overcome the fracture. According to this reading, the key to that unity lies in the rationalist principle of sufficient reason, which introduces the regulative demand that *is* and *ought* should coincide. Thus, an aspect largely ignored by mainstream Kantian interpretation comes to the fore: perplexity about evil is the impetus behind Kant's unification of theoretical and practical reason. The bafflement and threat of futility that overtake us when evil breaks the nexus of intelligibility drive the Kantian philosophical enterprise. For, as Rossi indicates, the most effective line of defense against evil is human solidarity, the promotion of which requires a drastic transformation of current social practices. Kant's philosophical ingenuity resides, then, in having channeled our metaphysical perplexity in the face of evil into productive practical uses. Critical philosophy is ultimately a kind of "anthropodicy," an immanent attempt at humanizing the world that makes transcendent flights into theodicy look outmoded and unwarranted.

Radical evil, "the foul stain of our species" (R 6: 38), it would seem, presents the most formidable obstacle against this project of human vindication. In "Kantian Moral Pessimism," however, Patrick Frierson shows how Kant's unflinching awareness of our moral deficiencies is not only compatible with moral progress, but also preferable to the anthropological optimism prevalent in contemporary moral theorizing. According to the latter, the main failings of human beings are explained by non-moral factors (knowledge, competence, social conditions, non-culpable negligence, etc.), which have little to do with "evil." This optimism pervades, for example, recent work in empirical social psychology (the situationism of Gilbert Harman and John Doris), and even the best normative ethics of Kantian extraction. As case in point, Frierson interprets central themes in Barbara Herman.

⁴ S. Neiman, *Evil in Modern Thought: An Alternative History of Philosophy* (Princeton University Press, 2002).

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)

Her rules of moral salience, analysis of non-moral motivation, and discussion of the impact of morality in our identity come under Frierson's fire. For they operate "under morally optimistic background assumptions" (p. 38). The problem is that these assumptions lead Herman to interpret our misdeeds in terms of factors for which we do not acknowledge full responsibility, and this interpretation legitimizes strategies Kant would consider self-ingratiating and self-deceptive. Although Kant's anthropological pessimism stymies these strategies, it does not let us fall into despair. On the contrary, Kant offers an inspiring vision of moral hope, "of endless progress [toward] complete conformity with the moral law" (KpV 5: 122). This hope, however, comes at a price: since the corruption of our moral character is radical, and at the same time it is our own fault, evil cannot be extirpated "through human forces" (R 6: 37) and requires the supernatural cooperation of God's enabling grace.

Kant's leap into transcendence is filled with tensions. In "Kant, the Bible, and the Recovery from Radical Evil," Gordon Michalson questions the feasibility of Kant's strategy to reduce the Bible to a rational/ethical core independent from theology. Michalson argues that Kant's appeal to the religious language of a "new man" and a "rebirth" to capture the temporal character of moral conversion does not work as it is supposed to, i.e., as a mere illustration of a self-standing moral argument. Rather, biblical references "serve as a *substitute* [for an argument nowhere to be found], as pictorial filler for a conceptual lacuna" (p. 58). Without this "filler," the moral community would lapse into apathy, for it would have no representation of what it is aspiring to. Yet, biblical references transcend the boundaries of applicability of Kantian concepts and are meant to account for a noumenal change that eludes rational explanation. Michalson detects, then, a fundamental aporia in Kant's *Religion*: on the one hand, it is necessary for us to *imagine* moral change in order to bring it about; yet, on the other, without violating the critical strictures, it is impossible to provide a conceptual account of such a change. Here is where biblical narrative comes to Kant's rescue: religious imagery "conveys the incommensurability between moral change and temporality while still offering language that helps us to represent the change" (p. 64). Although biblical language is not conceptual, it occupies a space whose void would otherwise be intolerable. "Biblical allusion thus becomes a kind of placeholder – an apparently indispensable placeholder – for the

narrative element that Kant's philosophical position requires but cannot provide" (p. 65). Michalson's analysis shows that religious narratives are not mere "parerga," as Kant used to believe, but have a function similar to the schematism in the first *Critique*. In both cases, something entirely rational (moral change, the categories) can be "represented" to the senses without erasing their respective boundaries.

II

Reduced to its bare essentials, Kant's conception of evil rests on three assumptions: (1) evil constitutes the underlying disposition of the human will (and hence is "radical"); (2) evil consists in the motivational primacy of the principle of self-love; and (3) there is a universal propensity to evil in all human beings, even the best.

All these assumptions are ripe for dispute. In "Kant's Moral Excluded Middle," Claudia Card argues that Kant's conception is flawed in at least two fundamental ways. First, Kant's theory of the will is "rigorist" and thus excludes all moral conditions that might be called intermediary, i.e., "neither good nor evil." Motivating Card's concern is the suspicion that the human will may not be a unitary, uniform, and internally consistent decision-making mechanism, as Kant presumed it to be. The best evidence we have to discover the nature of our will consists in the patterns of choice we observe over time. Here, Card notices, phenomena overwhelmingly point at the presence of conflicting volitional patterns, which suggest ambivalence and pluralism not the monolithic picture Kant favors. Furthermore, Card maintains that not all moral wrongs are evils: "culpability increases, other things equal, with increase in the harm the perpetrator is wrongfully willing to inflict" (p. 75). According to Card, Kant's harm-insensitivity sets him at odds with ordinary moral judgments: Kant's exclusive concern with culpability not only leads him to conflate serious and minor transgressions, but also to overlook the widespread phenomenon of having "moral scruples" and "making concessions" to morality, even among those who are committed to the principled pursuit of self-love. Kant can be spared from these blunders and remain true to himself, Card suggests, by incorporating a harm-sensitive dimension to his theory. "Radical harm," then, would complement Kant's "radical culpability," bringing radical evil in line with our ordinary judgments.

In “Evil Everywhere: The Ordinariness of Kantian Radical Evil,” Robert B. Louden mounts a sustained defense of Kant’s position against the most frequent objections in the literature. Most criticisms, Louden argues, rest on misunderstandings – once they are cleared away, the alleged shortcomings prove to be “in fact a strength” (p. 95). To begin with, Louden dismisses the objection of explanatory impotence, most thoroughly developed by Richard Bernstein.⁵ This criticism is off target: Kant never sets out to explain *why* human beings use freedom the way they do. Due to our epistemological limitations such explanations would be self-defeating: the source of free acts and the nature of our motives are inscrutable in principle. This does not mean, of course, that evil must be passed over in silence. Kant unambiguously identifies self-love as “the source of all evil” (R 6: 45). But, again, this identification seems naïve and disappointing to many interpreters. As H. Arendt famously argued, horrendous crimes cannot be explained “by comprehensible motives” such as “self-interest, greed, covetousness, resentment, lust for power, and cowardice.”⁶ All these motives fall under the rubric of self-love, and this principle seems too shallow to account for the totalitarian rendering of “all men . . . equally superfluous,” a crime that “breaks down all standards we know.” Although at one time Louden was sympathetic to this line of thought, he now maintains that self-love is a broad motivational notion and should not to be confused with selfishness. For Kant, the problem with self-love is that it refuses to recognize moral restrictions.⁷ Moral incorrigibility, not egotism or a trivial concern for happiness, is what makes self-love a candidate for “evil.” Thus interpreted, self-love is a motivational source capable of encompassing a variety of distinct types of desires and inclinations, and is even compatible with a great deal of unselfishness. It is not necessary, then, to invoke a diabolical will to account for egregious moral transgressions. Kant’s rejection of diabolical evil has nothing to do with the limitations of his moral

⁵ R. J. Bernstein, *Radical Evil: A Philosophical Interrogation* (Cambridge, MA: Polity, 2002).

⁶ Hannah Arendt, *The Origins of Totalitarianism*, new edition with added Prefaces (San Diego: Harcourt, 1994), p. 459.

⁷ Louden follows Andrews Reath here. See A. Reath, “Kant’s Theory of Moral Sensibility,” in *Agency and Autonomy in Kant’s Moral Theory* (Oxford: Clarendon Press, 2006), p. 16.

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)

Introduction

7

psychology, as John Silber used to argue.⁸ It rests on the grounds that moral accountability requires the capacity to consciously judge one's actions as being contrary to the moral law. The outright rejection of morality would turn the agent into a wanton, incapable of making moral discriminations, and thus unanswerable for the havoc she wreaks.

In "An Alternative Proof of the Universal Propensity to Evil," Pablo Muchnik develops an argument to justify the synthetic a priori character of Kant's claim "man is evil by nature." His strategy is to draw a systematic distinction between the seemingly identical concepts of "disposition" (*böse Gesinnung*) and "propensity" (*Hang zum Bösen*). While the notion of "disposition" indicates the fundamental moral outlook of an *individual agent*, the notion of "propensity" is meant to refer to the moral character of the *whole species*. The single appellative "evil," therefore, ranges over two different types of moral failure: an "evil disposition" is a failure to realize the *good* (i.e., to give duty motivational priority), whereas an "evil propensity" is a failure to realize the *highest good* (i.e., to engage in the collective project of shaping nature according to the demands of freedom). The correlation between units of moral analysis and types of obligation, Muchnik contends, clears the path for a philosophical justification of Kant's infamous claim: the attribution of radical evil to the species hinges on the same anthropological limitations that give rise to the doctrine of the highest good. According to this reading, Kant's proof is not really missing, as many interpreters have argued, but misplaced and buried where no one expects to find it, namely, in the Preface to the first edition of the *Religion*. Kant's coveted proof, Muchnik acknowledges, will probably disappoint the purists, since it falls short of the strict demonstrative standards of the first *Critique*. There is no denying it: the "transcendental" argument Kant advances in the *Religion* incorporates elements of his moral psychology arrived at by experience and is unabashedly "impure." Yet, it goes a long way to justify the subjective

⁸ J. R. Silber, "The Ethical Significance of Kant's *Religion*," in T. M. Grene and H. H. Hudson (eds.), *Religion within the Limits of Reason Alone* (La Salle: Open Court, 1934; 2nd edn, New York: Harper & Row, 1960), and "Kant at Auschwitz," in G. Funke and T. M. Seebohm (eds.), *Proceedings of the Sixth International Kant Congress* (Washington, D.C.: Center for Advanced Research in Phenomenology and University Press of America, 1991), pp. 177–211.

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)

necessity, universality, and a priori character of the propensity to evil. “[Its] *hybrid* nature ... is in line with the general thrust of the *Religion*, a book whose moral anthropology has also a quasi-transcendental ring, neither reducible to empirical observation nor totally severed from it” (p. 118). By striking a middle ground, Muchnik’s alternative proof is intended to solve “an unfortunate dilemma Kant poses to the interpreter: either to emphasize the widespread social/empirical dimensions of evil at the expense of its noumenal origin (the path Wood follows), or to stress its noumenal origin at the expense of its social/empirical dimension (Allison’s alternative)” (pp. 127–28).

III

Even if the reader were convinced by Kant’s controversial assumptions regarding rigorism, self-love, and the infamous claim that “all human beings are evil by nature,” the problem of how best to interpret evil still remains. In “Kant and the Intelligibility of Evil,” Allen Wood argues that a *sine qua non* for taking evil seriously is to regard it as “intelligible” – that is, as an objective phenomenon we have decisive reasons for *not* doing. But, if an evil action is one there are decisive reasons *not* to do, then evil is a species of motivated irrationality, a coherent description of which is notoriously difficult. According to Wood, Kant tackles this problem in two stages: first, he identifies “the fundamental maxim of evil,” which allows him to conceptualize “evil choices as following a highly general pattern” (p. 150); secondly, he interprets “this general pattern ... as fitting into human nature as it shows itself under the conditions in which human life has developed on earth” (ibid.). Wood calls these two explanatory stages “the maxim problem” and “the propensity problem,” respectively. We need the second, broader sense of intelligibility, because without understanding why evil is such a persistent feature of the human condition, we would not know how to struggle against it. This becomes clear if one relates the *Religion* with Kant’s essays on history, where he identifies radical evil with the dynamics of “unsocial sociability.” According to Wood, “*the human propensity to evil arises in the social condition, and develops along with the processes of cultivation and civilization that belong to it*” (p. 159). These processes bring about a situation of mutual dependency tied up with an anxiety “to gain worth in the opinion of others” (R 6: 37). Although originally a desire for equality, this anxiety gradually (though ineluctably, given the development of civilization)

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)*Introduction*

9

becomes a striving for ascendancy, i.e., “an unjust desire to acquire superiority for oneself over others [upon which] can be grafted the greatest vices of secret or open hostility to all whom we consider alien to us” (ibid.). Linking the moral excesses of individual and collective competitiveness with the development of social organization, Kant renders evil as intelligible as it can be. As a consequence, institutional arrangements become the battleground for moral progress, because it is at this level that the competitive tendencies associated with radical evil can be better controlled. The nub of Wood’s interpretation, then, is that evil is “a mechanism employed by natural purposiveness in developing our species’s predispositions in history” (p. 163).

In “Social Dimensions of Immanuel Kant’s Conception of Radical Evil,” Jeanine Grenberg finds three basic difficulties with Wood’s account: (1) it tends to undermine the individual’s responsibility and autonomy; (2) it obliterates the transcendental origin Kant attributes to the propensity to evil; and (3) it overlooks the fact that, unfortunately, evil takes many forms. Although Wood clearly is an individualist when it comes to moral responsibility, Grenberg finds a troubling ambivalence in the explanatory role he attributes to society in the genesis of evil. There is a trivial sense in which the presence of others provides a materially necessary condition for injuring them. But Wood, Grenberg contends, is claiming more than that: he endorses the Rousseauian view that in solitude the individual is good and tranquil, and it is people that “mutually corrupt each other’s disposition” (R 6: 93). Undoubtedly, the social setting provides the most notorious example of our competitive/comparative frame of mind. Yet, in the Kantian account, the propensity to evil must pre-exist our social engagements. Blaming others for my own moral corruption is a form of self-deception – a symptom of the inversion of the ethical order of priority, not an explanation of how it came about. Grenberg’s complaint, then, is that Wood confuses the cause with the symptom, and this confusion tends to dilute our individual responsibility. Furthermore, Grenberg takes issue with the problematic empirical status of “unsocial sociability,” the cornerstone of Wood’s interpretation: “reducing evil to a tendency in our interactions with other persons, Wood seems to have forgotten both that choice of this propensity is ‘prior to every use of freedom’ ... and that evil is a tendency to place concerns for self over ‘morality’ or ‘the moral law’ (R 6: 36), not simply over ‘others’” (pp. 178–79). To support this last point, Grenberg develops an account of the “social” in Kant, which she identifies with “shared purposes.”

Cambridge University Press

978-1-107-41473-0 - Kant's Anatomy of Evil

Edited by Sharon Anderson-Gold and Pablo Muchnik

Excerpt

[More information](#)

Not all moral transgressions can be reduced to this sphere: suicide, for instance, contravenes the duty of self-preservation (associated with the predisposition to animality), but does not necessarily undermine “shared purposes.” Grenberg’s point is that the possibilities for evil exceed the limits of the predisposition to humanity and the dynamics of “unsocial sociability.” Morality does not simply overlap with what we share with one another. Regrettably, evil has a polymorphic character and is irreducible to a single form.

A reply to this type of criticism can be found in the last section of Wood’s essay. There Wood argues that the social dynamics of evil are compatible with Kant’s commitment to transcendental freedom. Furthermore, to the extent that the propensity to evil is meant to elucidate “why we have a propensity to give the rationally weaker incentives of inclination or self-love priority over the rationally stronger incentives of morality” (p. 167), and that it is in the social condition that we come to value our status in the eyes of others more than our dignity as moral persons, Wood contends that the propensity to evil should not be limited to the violations of duties toward others, but also includes the condition for the possibility of violating duties to oneself. At the end of the day, in Wood’s reading, Kant’s appeal to the social condition “provides the necessary context for developing our radical propensity,” but does not entail that “society forces us to choose evil maxims, removing or diminishing our responsibility for these choices” (pp. 168–69). According to Kant, good or evil is always up to us, and those who blame society for their corrupt disposition are already “morally bankrupt” (p. 169).

IV

To give a taste of the relevance of Kant’s view to contemporary moral discussions, we conclude our *Anatomy of Evil* with reflections on genocide and moral reconstruction.

Because of its collective nature, extraordinary moral gravity and scope, genocide seems to mock our hopes for moral progress. Although no philosopher has championed the value of humanity more forcefully than Kant, genocide represents a form of “radical harm” of the type Claudia Card holds Kant is not prepared to accommodate. In “Kant, Radical Evil, and Crimes against Humanity,” Sharon Anderson-Gold challenges this conclusion. She argues that self-love, as it operates in individuals, is not limited to the “interests of the physical self”