

What Perfection Demands
An Irenaean Account of Kant on Radical Evil

Jacqueline Mariña

Critics of Kant's *Religion* have charged that the view of human nature there developed, and on which a great deal of the argument of the entirety of the book depends, is beset with a fundamental incoherence. On the one hand Kant claims that the radical evil corrupting the moral disposition is freely chosen: "This disposition . . . must be adopted through the free power of choice, for otherwise it could not be imputed." I will call this the *moral responsibility* thesis. On the other hand Kant claims, almost in the same breath, that radical evil is universal, attributable to all members of the human species ("there is no cause for excepting anyone from it," [R 6:25]). This is the *universality thesis*. It can be seen immediately that both claims seem to be at odds with one another: if we are truly free to choose our moral character, then prima facie we are not warranted in claiming that *all* human beings (all those that have existed, and those yet to be born) will have a morally corrupt disposition. Genuine freedom should at the very least imply that avoidance of the moral corruption of the disposition is *really possible*. This rules out the a priori universal imputation of radical evil to all human beings. Although it may just turn out that all human beings in fact happen to choose radical evil, nothing can be said about this choice in advance. At best, Kant is warranted in making only an empirical generalization, which is a far cry from the strong claim that no one is exempt from radical evil. As Philip Quinn has noted,

[S]ince the propensity to moral evil is a product of human freedom, it cannot be an essential element in human nature as is the predisposition to good. If moral evil is to be attributed to man as a species, it must be a contingent and accidental attribute of each member of the species . . . It seems clear enough that the prior probability of all humans choosing freely a morally evil supreme maxim must be quite low.¹

¹ P. Quinn, "Original Sin, Radical Evil, and Moral Identity," *Faith and Philosophy*, 2 (1984), 194. The problem has been noted by interpreters too numerous to cite. Henry Allison, for instance, notes that

Other contemporary interpreters (for instance, Michalson and Wolterstorff) have given up on the cogency of Kant's text as well, claiming that its incoherence is due to Kant's attempt to straddle two ultimately incompatible world views, that of the Enlightenment and Christendom: Kant wants to preserve both a Pelagian freedom and the notion of a fall and need for divine aid.²

In this chapter I will show that the incoherence many commentators have found in Kant's *Religion* is due to Augustinian assumptions about human evil that they are implicitly reading into the text.³ Eliminate the assumptions, and the inconsistencies evaporate: both theses, those of universality and moral responsibility, can be held together without contradiction. The Augustinian view must be replaced with what John Hick has dubbed an "Irenaean" account of human evil, which portrays the human being and his or her task in developmental terms. This developmental model is put forward by Kant in both the *Lectures in the Philosophy of Religion* and in his *Conjectural Beginnings of Human History*. In this chapter, I discuss both the Augustinian and Irenaean accounts of human evil and argue for the advantages of the developmental (Irenaean) account

an appeal to experience can at best "show that evil is widespread, not that there is a universal propensity to it." H. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), p. 154.

² For instance, in his book *Fallen Freedom: Kant on Radical Evil and Moral Regeneration*, Michalson argues that Kant's *Religion* contains a series of conceptual "wobbles" and "instabilities" that are "the inevitable result of a divided cultural inheritance" (p. xi); (Cambridge University Press, 1990); he presents a similar argument in his book *Kant and the Problem of Evil* (Blackwell, 1999). And in his chapter "Conundrums in Kant's Rational Religion," Nicholas P. Wolterstorff finds *Religion* so full of "conundrums," resulting from arguments pulling him in two different directions that his project is threatened with incoherence; in P. J. Rossi and M. Wreen (eds.), *Kant's Philosophy of Religion Reconsidered* (Bloomington: Indiana University Press, 1991).

³ There have been numerous interpretive strategies put forward to show that Kant is entitled to his universality thesis. Henry Allison, for example, treats Kant's claim as a synthetic a priori postulate and constructs a missing derivation for it (*Kant's Theory of Freedom*, p. 155ff). Other interpreters such as Sharon Anderson-Gold and Allen Wood argue that radical evil is "an empirical thesis," that cannot be reduced "to a mere inductive generalization. On this interpretation, radical evil would pertain to us insofar as we are social beings: the evil in our nature is closely bound up with our tendencies to compare ourselves with others and compete with them for self-worth" A. Wood, *Kant's Ethical Thought* (Cambridge: Cambridge University Press, 1999), p. 287; cf. S. Anderson-Gold, "God and Community: An Inquiry into the Religious Implications of the Highest Good," in *Kant's Philosophy of Religion Reconsidered* (Bloomington: Indiana University Press, 1991). There is something to be said for either approach, although something is still wanting in both. Allison is right in pointing out that the universality of radical evil is closely bound up with the fact that human beings do not have holy wills. His analysis falls short, however, in not recognizing that the key to making sense of Kant's text has to do with the difference between the disposition and the propensity, as I argue later. Moreover, Allison does not tie the lack of holiness up with Kant's developmental account. Both Wood and Anderson-Gold provide insights into the social pressures leading up to the universality of evil, but they do not couple their account with issues having to do with the moral development of the species.

in making philosophical sense of Kant's texts. I show that Kant indeed held such a view in both the *Lectures* and in the *Conjectural Beginnings*, and that he never abandoned the developmental model in *Religion*. Reading Kant's *Religion* through an "Irenaean" lens reframes the locus of the debate, illumines multiple elements of the text that previously remained obscure, and demonstrates why Kant had good reason to claim that we all begin in a condition of radical evil but must nevertheless assume responsibility for this.

I Two Models of Human Evil

There are two grand models of human evil and how it fits in with human destiny. The first model is the Augustinian. Here we have human beings, fully formed and in the presence of God, who inexplicably turn away from God and fall into evil. For our purposes, what is important about this model is the *inexplicability* and *seriousness* of the fall into evil. First, the fall is inexplicable. If these beings are fully mature, then presumably this implies they have an understanding of what is of true value. Insofar as they are in the presence of God, they are enjoying the beatific vision. But something tempts them away from that. How can this be? Second, the fall is serious. Given the original righteousness of these beings, who must have had a deep understanding of the value of that which they eschewed in turning away from the good, the fall is an extremely serious moral evil, so serious, in fact, that in some minds it justified the eternal damnation of fallen humanity.

The second model is associated with a minority view in Christian theology, namely that of St. Irenaeus (120–202 C.E.). In this model, creatures were created as morally and spiritually immature, and the fall was due to this immaturity. The fall is then like the mistake of a child that does not fully understand his or her identity, how the world really works, or what it is that has true worth. John Hick has pointed to the significance of this model in the development of his soul-making theodicy. In this model, "people were created as spiritually and morally immature creatures, at the beginning of a long process of further growth and development."⁴ God does not create creatures fully formed, but rather as simple and immature. Virtue and knowledge, the perfection of the creature, these are things the creature must attain for itself, that is, it must make itself into what it is intended to be through its own development of the seeds of

⁴ J. Hick, "Soul Making Theodicy," in S. T. Davis (ed.), *Encountering Evil* (Louisville, KY: Westminster John Knox Press, 1981), p. 318.

goodness implanted within it. Without such *self*-development, there is no genuine knowledge or understanding, for the individual has not really learned through its own experience and suffering of the consequences of good and evil. Without experience, the understanding of what has true worth cannot be truly *interiorized*. As Hick notes, “virtues which have been formed within the agent as a hard-won deposit of her own right decisions in situations of challenge and temptation, are intrinsically more valuable than virtues created within her ready made and without any effort on her part.”⁵ The possibility of the genuine interiorization of virtue and knowledge requires a *developmental* model of the human being. Here the human being makes mistakes and suffers their consequences. Had the creature been created perfect to begin with, the virtues would sit there, working lifelessly and autonomously, with no real knowledge or deep understanding of the consequences of good and evil infusing their being, for a genuine understanding of these consequences and their significance cannot be had except through experience.

The developmental model has the advantage that in it, the choice of evil is both *explicable* and not as catastrophic as on the Augustinian model. The choice of evil is explicable because the immature agent does not yet fully grasp the consequences of its choices. These only become clear to it over very long periods of time, as it experiences their working through. Moreover, once the difficult to attain knowledge of what has true value is in place, nothing can tempt the individual to regress. Perfection is hard won and irreversible. (In the Augustinian model, on the other hand, there is no guarantee that once the soul has regained its original perfection, it will not fall again.) The choice of evil is furthermore not as catastrophic as on the Augustinian model. Because all beings begin in a state of simplicity and ignorance, bad choices are not infinite offenses deserving infinite punishment, for evil is not the turning away from the good of the fully formed finite creature possessing original righteousness and true understanding. Rather, bad choices are expressive of the immature agent’s level of knowledge and moral development.

II Kant’s Developmental Model

In both the *Lectures on the Philosophical Doctrine of Religion* and in the *Conjectural Beginnings of Human History* (1786) Kant offers us a developmental model of the human being. In *Conjectural Beginnings* we are

⁵ *Ibid.*, p. 320.

told that the vocation of the species is “nothing but a *progressing* toward perfection” (8:115), and in the *Lectures* Kant notes that “among the many creatures, the human being is the only one who has to work for his perfections and for the goodness of his character, producing them from within himself” (28:1077). We begin from a condition of imperfection; our starting point is our animal nature, which it is our task to transcend.

Both texts ingeniously reinterpret the story of the fall. In *Conjectural Beginnings* we are told that the departure from paradise is “nothing other than the transition from the crudity of a merely animal nature into humanity, from the go-cart of instinct to the guidance of reason – in a word, from the guardianship of nature into the condition of freedom” (8:115). Evil emerges when reason comes on the scene. So long as human beings live only in the moment, guided by instinct alone, they are in paradise. Because at this point in their development reason does not yet exist in them and thereby makes no demands, they also incur no guilt. In paradise the human being “lives undisturbed, led by his instincts, until finally he feels his humanity, and in order to prove himself, he falls. Here he no longer *is* an animal, but he has become an animal” (*Lectures*, 20:1078). That is, once reason begins to stir, the person no longer simply *is* what she finds herself to be. Reason demands that she must *make herself* into what she is to become. Because reason represents a moral demand and hence a choice, she becomes responsible for persisting in her condition of animality insofar as she does so, for she has chosen to identify with her animal needs. In this way, she has *made herself* into an animal.

Strikingly, Kant concludes that “*the first development of our reason toward the good is the origin of evil.*” Once the subject is able to compare its present state with what it should become, it becomes responsible for its condition. And once reason enters the picture the subject is split with itself; it can no longer be satisfied to rest in the animal passions that it has, but understands the demand that it transcend this animal nature and so becomes unhappy with itself:

Before reason awoke, there was neither command nor prohibition and hence no transgression; but when reason began its business and, weak as it is, got into a scuffle with animality in its whole strength, then there had to arise ills and, what is worse, with more cultivated reason, vices, which were entirely alien to the condition of ignorance and hence of innocence. The first step out of this condition, therefore was on the moral side a *fall*; on the physical side, a multitude of ills of life hitherto unknown were the consequence of this fall, hence punishment. (8:115).

When the human being makes this beginning, he first uses his reason in the service of instinct; finally he develops it *for its own sake*. Hence he finds evil *first* when his reason has developed itself far enough that he recognizes his obligations. St. Paul says that sin follows upon the law. (28:1079)

Before the birth of reason, the human being is sunk in sensuousness, concerned only with the gratification of its immediate desires. Because it has direct access only to its own conscious experience and little power of the imagination, it is concerned only with itself. But reason begins to change all of this, expanding the individual's horizon in multiple ways. Reason's entrance is tied to the power of the imagination. Through the use of both reason and imagination the human being begins to make comparisons, notices causes and effects, understands consequences, and thereby foresees possible futures. Furthermore, she begins to notice the others: they too, are rational beings having a first-person experience, and must be considered; their existence places demands on her, for they cannot be considered as mere means (8:III–II5). Reason, however, does not first come on the scene in its full maturity and strength, but only in its infancy, its powers weak. At first it is always the worsted party in its "scuffle with animality in its whole strength"; it is to be expected that its first exercises will be almost fully bungled, for along with reason come fresh desires due to the expanded powers of the imagination, and hence a new array of possible mistakes. There is just too much to manage. As such the entrance of reason brings with it "only false steps and foolishness" (28:1077).

The two fundamental practical demands for finite rational beings are the cultivation of talents and benevolence of the will (28:1077).⁶ These are *demands*, for the finite rational agent is a being of needs. These needs clamor for attention and at first fully obscure the demands of reason. The finite agent's desire for immediate gratification will tempt it toward indolence and impede its capacity to understand the standpoint of the other. For instance, the needs of the individual will limit benevolence, which is "an immediate well-pleasement with the welfare of others." Pure benevolence "is nowhere to be found. For every creature has needs which limit its inclination to make others happy, or its *de facto* ability to exercise these inclinations in such a way as to have no regard at all for its own welfare" (28:1076). Only God is capable of a pure and complete benevolence, one limited only by justice and holiness.

⁶ These two comprise our duties to self and to others, discussed at length in major works such as the *Groundwork* and the *Metaphysics of Morals*. Our benevolence toward particular individuals must, of course, always be informed by the larger demands of morality.

Because of our needs and inclinations, "the human being can *never* be *holy*, but of course [he can] be virtuous. For virtue consists precisely in *self-overcoming*" (28:1075). This is key to Kant's anthropology and its influence on his moral theory. Because the human race begins from an immature state, there is no question that its individual members can discharge the demands of reason when the race is in its infancy. We can expect that the first exercises of reason will always end in foolishness. Yet this immature beginning is something that must be presupposed if the demand that the human being *develop itself* is to be met, that is, if its virtues are to be hard won. And only if its virtues are hard won will the human being fully *understand* the good. As such, the task of the human being is always a *self-overcoming*, or a transcending of the self. It must overcome both the temptation to immediate gratification and its immediate concern with itself. But the idea of a *self-overcoming*, of course, presupposes that the individual begins with itself and the limitation of its self-concern. In its infancy, reason can only begin to coax individuals out of their solipsistic ways. Many trials, false steps, and entanglements await both the individual and the race, although "the whole is *someday to win through to a glorious outcome*, though perhaps only after enduring many punishments for their deviation" (28:1079). There are, as such, no magical solutions to the overcoming of evil, which is due to limitations in development. All lessons must be appropriated, that is, made one's own, and they can be appropriated only through experience, that is, the repeated trials of self-overcoming through the use of reason. Only through appropriation is genuine understanding possible. But once the human being acquires genuine understanding and develops fully, nothing can tempt her to do evil: "When the human being has finally developed himself completely, evil will cease of itself," (28:1079), for evil is only "*incompleteness in the development of the germ of the good*" (28:1078).

III Radical Evil and Self-Overcoming in *Religion*

From this perspective Kant's claim regarding the universality of human evil in *Religion* would certainly be understandable. It is due to the infancy of the race and the corresponding state of the individuals who are members of it. They begin in radical evil not because they have "fallen" from a state of perfection, but rather because of the relative newness of their capacity to reason and the clumsiness with which they begin its exercise. Yet in *Religion* Kant's thinking about the nature of evil is complicated in significant ways. This transformation does not, however, vitiate his

developmental analysis of the human being. In *Religion* Kant does not all of a sudden adopt an Augustinian view of original sin. However, the fundamental problem of the *Religion*, namely that of radical evil, significantly complicates the “Irenaeian” picture sketched out earlier, particularly in regard to the question of the moral responsibility of developing rational agents. As I show later, Kant’s analysis of human evil as radical is an implication of his claim that pure reason can be practical. Because pure reason is practical, we are responsible for our misdeeds: if there is evil, it is because it has been chosen over the incentive of morality. But what might this mean for a developing being? To what degree can the quandaries of *Religion* be attributed to Kant’s attempt to retain the developmental model of the species⁷ and the individuals that comprise it (the propensity to evil is “in all cases somehow entwined with humanity itself” [6:32]), while at the same time providing an account of the imputability of our missteps, indeed, even of the initial corruption that gave rise to them? Only if those missteps are imputable can the injunction to become *better* really make sense.

At the beginning of the second book of *Religion* Kant reflects that “to become a morally good human being it is not simply enough to let the germ of the good which lies in our species develop unhindered; there is in us an active and opposing call of evil which is also to be combated” (6:57). Here we find Kant deepening, and perhaps even revising, his earlier claim that evil is “only incompleteness in the development of the germ of the good.” Although evil may still be an incompleteness in development, it is not *simply* this, for along with this incompleteness is always found an *active resistance* to the demand toward self-transcendence or self-overcoming. This resistance originates in the free *Willkür*, which, as will be shown later, continuously values *itself* – not only its continued survival, but also its capacity for *self-assertion*, over moral demands. Kant stresses that once reason comes on the scene, the moral demand is not only clear, but it also announces itself as the “highest incentive” (6:26, note). The human being is not only conscious of moral demands, but also is

aware that they have the most value. Yet even with this awareness the human being fails to live up to what it is called to do. This failure can only be accounted for by the positing of what Kant calls radical evil in human nature.

However, whereas Kant significantly complicates the developmental model of the *Lectures* and *Conjectural Beginnings* in *Religion*, a careful reading of the latter text reveals him still committed to the Irenaeian view. Goodness is something that must be *achieved*. Echoing many of the ideas in the *Lectures*, Kant notes that “the human being must make or have made himself into whatever he is or should become in a moral sense, good or evil” (6:44). And commenting on the creation of the human being as good, he is careful to explain that this means not that she was originally created *already good* (as in the Augustinian picture), but that she is created *for the good*, that is, her goodness is set to her as a task: “the human being is not thereby good as such, but he brings it about that he becomes either good or evil” (6:45). She must perfect herself through the development of the predispositions that lie in her as germ. In *Religion* “The Fall” is not a descent from a condition of original perfection, where the human being had already mastered and exercised goodness, and then inexplicably fell away from that condition. It is rather an account, from a purely rational perspective, of how it is possible that free beings *predisposed* toward the good nevertheless fail in its exercise. Here Kant argues that this failure is not due to immaturity alone. Something much more sinister, a principle of willful self-assertion, is also at work. We do not struggle against the inclinations alone, but “against principalities and powers” (6:59), that is, against evil principles we have ourselves adopted. As such, the work of *self-overcoming* is just that much more difficult. Yet, it may very well be that this *self-overcoming* could not be set for us as a task unless we first began with those evil principles. This would be the insight that Hegel, following the clues set out by Kant, would develop.

The significance of Kant’s concept of radical evil cannot be appreciated without an analysis of what he calls the fundamental disposition. Throughout *Religion*, he uses “nature” in two distinct ways. First, he refers to the nature of the species, and second, he speaks of the nature of the individual. In both cases, he is at pains to show that we can, indeed, speak of the nature of both the species and the individuals comprising it without compromising the freedom of the *Willkür* (the power of choice) that must be posited if the imputation of actions is to be possible. Each moral agent has a fundamental disposition, which, although freely chosen, functions very much like a person’s character or nature. This disposition grounds any

⁷ As Firestone and Jacobs have stressed in *In Defense of Kant’s Religion*, Kant arrives at the universality of radical evil by attributing it to the species as a whole, and by implication to each member of it. As Kant notes, “by the human being of whom we say that he is good or evil by nature we are entitled to understand not individuals (for otherwise one human being could be assumed to be good, and another evil, by nature) but the whole species” (R 6:25). That Kant’s claim concerning the universality of evil is restricted to the human species is noted by Robert Louden in his essay, “Evil Everywhere: The Ordinarity of Kantian Radical Evil,” in S. Anderson-Gold and P. Muchnik (eds.), *Kant’s Anatomy of Evil* (Cambridge University Press, 2010), pp. 93–111; cf. R. Louden, *Kant’s Impure Ethics: From Rational Beings to Human Beings* (New York: Oxford University Press, 2000), p. 132.

choice of principle, decision, or action that occurs in time. It is “the first ground of the adoption of our maxims, which must itself again lie in the free power of choice” (6:22).

Importantly, Kant notes that the disposition “cannot be any fact [factum, i.e., something done] possibly given in experience.” There are two important reasons for this. First, when inquiring into grounds, we must distinguish between *reasons* and *causes*. Because anything given in experience must be taken as an *event*, its determining grounds must be referred to preceding states, that is, causes occurring in time. But because any natural cause must itself have a cause, when seeking grounds in the order of causes, the ground of the individual’s agency would have to be referred back to causes ultimately lying outside the will itself. This would conflict with the imputability of actions and moral responsibility. Hence, we must seek this ground “merely in the representation of reason” (6:39). Second, were we to inquire into the ground of any maxim chosen in time in terms of reasons, its ground could only be found in another freely adopted maxim. To avoid an endless regress, we must posit a fundamental maxim grounding all other choices that manifest themselves in time (6:20). As I discuss later on, the choice of this fundamental maxim affects the *propensity* through which all her other choices can be accounted for. When she makes a choice, we can inquire why things *mattered* to her in such a way that she made that choice. Certain things will have impressed themselves on her as more important than others. Kant’s point is that when it comes to moral concerns, *why* things matter to people in the ways that they do must ultimately be understood in terms of a freely chosen basic principle expressive of the complete *orientation* of the individual. Although it may be that in their daily experience persons may seem to simply *find* themselves caring about some things more than others, how things matter to people as they go about their lives must ultimately be grounded in a fundamental choice, one hidden in the depths of the soul, coloring both the affective and imaginative powers of the individual as she acts in the world. Such actions are to be understood as mere *manifestations* of the fundamental orientation constituting the person’s character, an orientation that has been, however, chosen and is not the mere result of what has occurred previously in the world.

Grounding as it does all subsequent choices, the disposition is singular. This is Kant’s rigorism thesis. Its singularity is due to the fact that the individual’s character is always already *oriented* in a particular way, and this *orientation*, grounding as it does all other choices, cannot be both good and bad at one and the same time. The individual never begins from a neutral

standpoint, so that at one moment she can turn toward good and the next toward evil. The impossibility of such a neutral ground is implied by the very nature of good and evil; the choice of each has a particular kind of impulse and *directedness* to it, so that, for instance, the turning toward evil does not implicate one in a single evil deed but in a whole way of life. That moral character is always already *oriented* is implied by the fact that the moral law is an incentive: because it is an incentive, failure to act in accordance with it must be due to an active *resistance*, which means that the individual has actively oriented herself *against* the good, and this orientation is never operative in just a single deed. This is the significance of Kant’s note at the beginning of book one, in which he argues that the “not good” is “a positive antagonistic to the good = -a.” The reason for this is because the law is an incentive, “the lack of agreement of the power of choice with it (= o) is possible only as the consequence of a real and opposite determination of the power of choice, i.e., of a resistance on its part, = -a, or again, it is only possible through an evil power of choice” (6:23).

Two more notions are crucial to Kant’s exposition of the disposition. The first is Kant’s incorporation thesis and the second his exposition of the character of good and evil in the individual. According to the incorporation thesis, something’s being an incentive is necessary, but not sufficient, to its being the ground of choice of an action. For that something more is required, and that is the free choice of the individual to recognize or reject the incentive. Hence an individual may in fact have powerful passions and desires in accordance with which a particular course of action is deemed quite attractive, but it is nevertheless up to her to reflect and consider whether in fact such desires are *worth* acting on. In affirming their worth she incorporates them into her maxim, that is, she chooses them as the sorts of desires *she* wants to have. This incorporation is done through the power of free choice (6:24). Each individual will have many desires, but she must choose which ones she is to ally herself with, and in such a way *make herself* into what she is to become.

Moral goodness has to do with what one cares about and how one goes about doing so. What is taken as most important, and what as less? There are many kinds of things that we must care about, but the question is, which kind of thing is valued more? In virtue of being human, we all have drives having to do with our embodiment. We also quite naturally have desires stemming from our wanting to be recognized by others. Kant emphasizes that the objects of these desires have merely a subjective worth. Their value derives from my subjectivity and individuality, and their satisfaction is tied up with self-love and the desire for happiness. In and

of themselves, both kinds of desires are necessary and good; Kant considers them *predispositions* determining human nature (6:26). The most important predisposition in us, however, is the predisposition to personality, “the susceptibility to respect the moral law as of itself a sufficient incentive to the power of choice” (6:27). The moral law alone is objectively valuable. Because each predisposition has its corresponding incentives, and all must be incorporated into one’s maxims, the question of whether the individual is good or evil must have to do with the order of importance that she assigns to each incentive, that is, “which of the two [self-love or morality] he makes the condition of the other” (6:36). Given the distinct nature of the *kinds* of desires associated with each predisposition, the problem may arise that satisfying one set of desires will conflict with satisfying the other. Aims can pull apart, and one can be driven in different directions. It is in this situation where human aims pull apart that the ordering of what we care about becomes most important and can most clearly be seen.

Conflicts between required human aims arise for two sorts of reasons. The first kind of reason has to do with a kind of incontinence. Here we have desires of a particular sort getting too large, too strange, or out of control and consequently interfering with other aims key to our existence as human beings. When this happens to the desires stemming from our animal nature, we get “vices of the savagery of nature” and “the *bestial vices of gluttony, lust, and wild lawlessness* (in relation to other human beings).” And when this happens to our desire for recognition, we find “an unjust desire to acquire superiority for oneself over others.” Hence “the greatest of secret or open hostility to all whom we consider alien to us” can grow out of this jealousy and rivalry (6:27). In both cases, what were at first natural desires are transmuted to vices, namely desires that, in and of themselves, conflict with the demands of morality. They grow so large that their satisfaction requires a kind of self-assertion over against the moral law; or more accurately, it is in and through the act of self-assertion that natural desires having to do with embodiment or recognition are deformed, for here the desires are put in the service of something other than their natural aim. Here *my subjective desires* take center stage, and I think of them as sufficient reasons not only for my action, but for your action, too. This is what Kant calls *self-conceit* in the second *Critique*, the attitude fundamentally at odds with the moral demand.⁸ The first case,

⁸ As Kant notes, “This propensity to make oneself as having subjective determining grounds of choice into the objective determining ground of the will in general can be called *self-love*, and if self-love makes itself lawgiving and the unconditional practical principle, it can be called *self-conceit*” (5:74).

then, in which the demands of morality can pull away from the desires stemming from embodiment and social recognition are due to this kind of self-conceit, where I consider my desires, whatever they may be, more important than treating others as ends in themselves.⁹

There is another case, however, in which moral aims can pull away from social and animal needs and require a decision, a ranking of values. This is true, for instance, of the just person in Plato’s *Republic*, who, due to her choice of justice for its own sake, winds up with the worst reputation, her eyes burned out and ultimately impaled. Kant brings up his own case in the second *Critique*: here we have an individual in an impossible political situation required by the king to give false witness against an innocent man on pain of death. The really good person chooses what morality demands for its own sake, and not with an eye to other aims. And this choice – the moral ordering of what is important – can be tested in precisely those situations of trial, where the choice must indeed be made to do the right thing, even when a great deal of personal sacrifice – even the ultimate sacrifice of one’s life – is involved. In this case all one’s personal aims are sacrificed for the sake of morality. This situation is very different than the one given earlier, for an individual may indeed be quite continent and respectful of others, but may yet not have endured, or be capable of enduring, this kind of test.

The two cases requiring a self-conscious choice between aims are quite different, but in the end Kant argues that only if the person is capable of surviving the latter test is she really a *moral* individual. This shows that *what perfection demands is indeed great*; it does not just involve the achievement of an accidentally correct ordering of desires, for this may occur because the individual has never really been tested. One may have

For an excellent account of Kant on self-conceit, see A. Reath, “Kant’s Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination,” in *Agency and Autonomy in Kant’s Moral Theory* (Oxford: Clarendon Press, 2006), pp. 8–32. There he notes that self-conceit is the “tendency to treat oneself or one’s inclinations as providing reasons for the actions of *others*, or to take one’s desires as sources of value to which they should defer . . . self-conceit would produce a form of first person egoism, in which I act as though *my* inclinations could provide laws for the conduct of *others*: it expresses a desire that they serve or defer to my interests.” (p. 15)

⁹ Self-conceit is certainly a kind of self-assertion over against what Hegel would call “the universal.” It is therefore an extraordinarily serious matter. Although Kant rejects the notion that the human being can be diabolical, that is, can choose evil for its own sake, he certainly recognizes the kind of motive worthy of Milton’s Lucifer, who would “rather reign in hell than serve in heaven.” In this case, evil is not chosen for its own sake, but the self and its merely subjective desires are made the condition of the worth of anything else. Here the self puts itself in center stage and attempts to usurp the place of God, or the moral law itself. For an excellent discussion of this problem containing a good review of the secondary literature on this point, see Loudon, “Evil Everywhere,” pp. 103–108.

found oneself among a very congenial social group of moral individuals, in which what is required of one to have nice things and achieve recognition also happens to dovetail with what morality demands. So in this case one can be quite continent and respectful of others, showing no signs of self-conceit. But how can one possibly know what one really considers most important here? What perfection demands is the willingness to sacrifice the self completely to the demands of morality. And not just once, but on a daily basis. Only here do we have *self-overcoming* in its fullest sense. Most of us find ourselves in *some* situations of trial, sometimes quite hidden, in which hard choices must be made. But to what degree should it be expected that the morally developing individual *always* passes these tests? Is failure to always pass ultimate moral tests the same as *self-conceit*?

The difficulty in answering this question – the complicated character of moral demands and our possibilities of responding to them – is recognized by Kant in his discussion of both the propensity and the three degrees of radical evil. Although the disposition is a fundamental maxim (subjective practical principle), the propensity, which is influenced by the fundamental choice, grounds the individual's emotional and affective life. Insofar as the propensity follows from the fundamental choice it can be considered a state grounding other choices: it is the fundamental orientation that results from choice. However, it is not simply the *result* of the choice of fundamental maxim, but it must be actively maintained in the order of reasons, and in this sense is a deed. However, the propensity is not a deed in the sense of a phenomenally recognizable act. In this regard it is “the formal ground of every deed” given in phenomenal experience (6:31). Kant notes that the “good or the evil heart” arises from the propensity (6:29); hence it has to do with how things can matter to us, and the kinds of importance we assign to our affairs. The propensity conditions the soul at its very root, at the common ground of sensibility and understanding. Because the individual has already made a grounding choice – too often unexamined – about what she takes to be most important, she will simply *find* the objects of her daily concern mattering to her in one way or another. The propensity is so deeply ensconced in the soul that it preexists experience and hence actual incentives. In fact, a person may not have certain inclinations only because they have not yet been awoken. As such, it may be quite accidental that a person is not subject to certain vices, for she may simply not have encountered the conditions that would have brought them out. The subjective ground for them is, however, still there (6:29). And because at the level of her inner phenomenal life she simply finds things she encounters mattering to her in a certain way, this has enormous effects on her imaginative

capacities, on what characteristics of a situation she will identify as salient, on what, for her, *counts*, and hence, on which moral demand she considers to be at play. In other words, the propensity conditions how an individual understands and interprets her situation and its demands. Consequently, it affects the capacity for moral judgment. This does not mean it vitiates consciousness of the demand of the moral law itself; that is always present and cannot be expunged. But it will color the individual's understanding of how to subsume particular cases under moral demands, and even the individual's self-understanding. Here the capacity for self-deception is indeed great. This is why the propensity is impossible to root out. The very capacity for moral understanding has *itself* been compromised by the propensity to evil. But it is this capacity that is first required to even *identify* it. The propensity to evil is therefore always opaque to itself, and once acquired, seems to be inextirpable.

Although the propensity, like the disposition, is singular, there are different degrees of its corruption. In the first, and mildest, case of radical evil, the self is very clear about the intensity, depth, and scope of the moral demand, wills to comply with it, but recognizes its inability to do so. Perfection demands a complete self-overcoming to which the heart may not be adequate. This is the case of frailty: the fundamental disposition is a good one; the individual has in fact incorporated the good into the maxim of its power of choice. Yet, here we have “the weakness of the human heart in complying with the adopted maxims” (6:29); the individual's affective and imaginative powers are not adequate to the exercise of the good disposition that has been chosen. It is important to note that in this case the propensity does not result from the disposition.¹⁰ Rather, the propensity to evil is due to immaturity. Here the fundamental disposition is good; the individual is certainly headed in the right direction, but its powers are not equal to the task and need to be further developed and strengthened.¹¹

¹⁰ Pablo Muchnik rightly points out that most interpreters have not been careful to recognize the difference between the propensity and the disposition. This difference, however, is key to making sense of Kant's text: “the notions in question refer to two different units of moral analysis: the *Gesinnung* indicates the fundamental moral out-look of an *individual agent*; the propensity, the moral character imputable to the *whole human species*. Overlooking the logical independence of these analytic units gives the impression that Kant's talk of a universal propensity to evil is inconsistent with his commitment to freedom” (p. 117), see P. Muchnik, “An Alternative Proof of the Universal Propensity to Evil,” in S. Anderson-Gold and P. Muchnik (eds.) *Kant's Anatomy of Evil* (Cambridge: Cambridge University Press, 2010), pp. 116–43. Although I do not agree with all points of Muchnik's analysis, he is certainly correct to point out that the propensity and the fundamental disposition, although closely tied, are two different things.

¹¹ Kant's recognition that the fundamental disposition can be good, whereas the propensity may still incline to evil, reveals the consistency of Kant's claim that even if the fundamental disposition is good,

For instance, the self has access only to its own first-person experience; at best it can only imagine the first-person experience of others. And not only are its powers of imagination weak, but it may not really *want* to think about the first-person experience of the other, either, for it is already overwhelmed with its own cares. This is certainly an impediment to the recognition that the other counts just as much as oneself, absolutely key to self-overcoming. For the heart to be adequate to the moral demand, it must widen the circle of its concern, of what matters to it, and hence how it understands and interprets and works through the moral demand. Even when the disposition is fundamentally good, all of these powers stand in need of development for moral goodness to be achieved.

The impurity of the human heart is at work in the second degree of the propensity to evil (6:29). Here we have motives that are not purely moral; a good bit of self-deception may be involved. In this case the individual chooses to remain unclear about the real costs of moral commitment and strives, whenever possible, to make both incentives, that of happiness and morality, work together. The individual makes prudent and easy choices. She avoids, or perhaps chooses not to recognize, situations that would require her to choose between comfort and the moral demand, all the while counting herself an honest and outstanding person. She can be blind to what the existence of others really means for her, especially when this might mean too much, that is, when the moral claim is too great. She does not want to see what this demand might imply, in terms of time, in terms of effort, or in terms of the obligation to stake her comfort to right some wrong or to help another. In some sense, she aims at fulfilling the moral demand, but she takes this demand to be an easy one. She wants to believe that the aims of morality and happiness can never pull apart, or don't do so, at least in most cases, so that in aiming at her self-interest in a clever way she can be moral, too. However, this attempt to not to have to choose between the aims of self-love and the moral demand amounts to a de facto choice of comfort as to what is most important, and here morality is the worsted party.

Kant calls the third degree of radical evil “the perversity of the human heart,” for here the ethical order of the incentives is reversed: the individual

there is still a need for divine aid. Hence Kant grants that “some supernatural cooperation is also needed to becoming good or better.” He notes, however that “the human being must make himself antecedently worthy of receiving” divine aid, that is, the person “must accept this help (which is no small matter, i.e., he must incorporate this positive increase of force into his maxim” (6:44). For a defense of Kant's claim that for divine aid to be effective it must be incorporated into one's maxim, see my article, “Kant on Grace: A Reply to His Critics,” *Religious Studies*, 33 (1997), pp. 379–400.

chooses to value the satisfaction of her own desires over the moral demand. What this ultimately means is that she disregards that others are ends in themselves, too, and that they cannot be used or ignored for her to fulfill her desires, whatever those may be. The third degree of radical evil thus involves full-blown self-conceit, a kind of assertion of the self over universal demands. Here desires having merely subjective validity are treated as if they were sufficient reasons for the actions of others. Others are used as mere means, and the legitimate claims and needs of others are simply ignored. Kant claims that in the first two cases of radical evil we have an unintentional guilt; in the last case, however, one treats oneself as both author and end (*telos*) of the law. This results in “deliberate guilt (*dolus*), and [it] is characterized by a certain *perfidy* on the part of the human heart” (6:38).

Kant's discussion of the origin of the evil heart reveals much about the relation between immaturity (frailty) and the reversal of incentives in genuine human evil. He notes:

An evil heart can coexist with a will which in the abstract is good. Its origin is the frailty of human nature, in not being strong enough to comply with its adopted principles, coupled with its dishonesty in not screening incentives (even those of well-intentioned actions) in accordance with the moral guide, and hence at the end, if it comes to this, in seeing only to the conformity of these incentives to the law, not to whether they have been derived from the latter itself, i.e., from it as the sole incentive (6:37).

This remark shows why Kant thought that the propensity is “woven into human nature,” while at the same time being bound up with freedom. It must be stressed that according to Kant's discussion of the first degree of radical evil, the *frailty* of the human heart can certainly coexist with a good fundamental disposition. As such, we can understand how the propensity can be universal while coexisting with the freedom of the *Willkür*; frailty does not, of itself, bring an evil disposition in its train. It results from lack of strength and inner development, but does not imply full-blown moral evil, namely, a reversal in the order of incentives. It can, however, still be considered a propensity to evil, insofar as it inclines but does not necessitate one to impurity or perversity. Impurity results from the attempt to overcome the distance between one's frailty and the moral demand through self-deception: one soothes oneself into believing that moral demands can be met through their coincidence with the satisfaction of other incentives. Finally, in the last case we have full-blown evil, in which the individual does “not trouble itself on account of its disposition but rather considers itself justified before the law” (6:38). Kant seems to suggest

that here we have a kind of Pharisaism, “the fancy that they deserve not to feel guilty of such transgressions as they see others burdened with” (6:38). This self-righteous attitude fits well with self-conceit. It puts the law in the *service* of the self and its merely subjective aims. The law is no longer a guide urging the overcoming of self-concern through a genuine encounter with others. It becomes, instead, a slave to self-assertion. It is that through which the self justifies itself as superior and having a special right over others. The law is used to judge them and find them wanting, thereby lacking worth and moral considerability. In the *Vigilantius* lectures, Kant describes this depravity as “a wickedness in the way of thinking about others” VE 27:691.¹²

The propensity to evil first appears as frailty. Frailty, however, does not imply a reversal of the moral incentives. This only occurs when the self deceives itself about its weakness and foregoes humility before the law, the only attitude appropriate to its character. In so doing the self finally puts itself in a position where it lacks compassion for the others, who are equally just as frail. The fact that the first degree of radical evil does not imply a reversal of the moral incentives shows that the propensity can indeed be universal, “woven into human nature,” without thereby necessitating full-blown moral evil. We can understand its being woven into human nature in terms of the infancy of the species, so that every human being, even the best, will be affected by the propensity to evil. When confronted with the enormity of the moral task, the human conditions of frailty and finitude tempt to both impurity and perversity without thereby necessitating them. Kant’s position in *Religion* complicates his developmental account. It does not, however, replace it with an Augustinian one. Moreover, this developmental picture goes a long way toward demonstrating why the propensity is indeed universal, even while preserving the freedom of the *Willkür*.

¹² A. Wood mentions this phrase to highlight the social content of evil, in Wood, *Kant’s Ethical Thought*, p. 288.

Atonement and Grace in Kant *Some Reflections*

Keith Yandell

What is the proper translation of Kant’s major work on religion – *Religion within the Limits of Reason Alone* or *Religion within the Limits of Mere Reason*? From one perspective, what follows is an answer to that question. From another, it is a query into what account Kant gives of atonement. The tasks are related.

I Introduction

Sometimes it is permitted that one take a broad look at a philosophical system, or part thereof, without going into details. This is particularly so for an article or a chapter on a topic regarding which there is systematic exegetical and philosophical dispute. Kant’s view of grace fits this description. Two authors who have presented differing readings of the philosophical situation express their views as follows.

One side of the dispute is nicely put by Nicholas Wolterstorff:

If we have a moral claim on someone’s doing something then for that person to do that is not for the person to act graciously but to grant what is due to us; it is to act justly, not graciously.¹

What follows is that Kant cannot, consistently with his own moral theory, suppose that a just God must forgive. This violates the “moral order of rights and obligations” for which Kant argues and leaves Kant without the logical space for gracious forgiveness. In a sense, this is correct, as we will see. Kant, after all, affirms categorically that justice is remunerative, not gift-making.

On another side is Jacqueline Mariña, who holds that he does have a doctrine of grace:

¹ N. Wolterstorff, “Conundrums in Kant’s Rational Religion,” in Philip Rossi and Michael Wreen (eds.), *Kant’s Philosophy of Religion Reconsidered* (Indiana University Press, 1991), p. 44.

KANT AND THE QUESTION OF THEOLOGY

EDITED BY

CHRIS L. FIRESTONE
NATHAN A. JACOBS
JAMES H. JOINER

 **CAMBRIDGE**
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India
79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.
It furthers the University's mission by disseminating knowledge in the pursuit of
education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107116818
DOI: 10.1017/9781316337813

© Cambridge University Press 2017

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2017

Printed in the United Kingdom by Clays, St Ives plc

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

NAMES: Firestone, Chris L., 1957- editor.

TITLE: Kant and the question of theology / edited by Chris L. Firestone,
Nathan A. Jacobs, James H. Joiner.

DESCRIPTION: 1 [edition]. | New York : Cambridge University Press, 2017. |
Includes bibliographical references and index.

IDENTIFIERS: LCCN 2017008317 | ISBN 9781107116818 (Hardback)

SUBJECTS: LCSH: Kant, Immanuel, 1724-1804. | Religion—Philosophy. |
Philosophical theology. | Theology.

CLASSIFICATION: LCC B2799.R4 K357 2017 | DDC 210.92—dc23

LC record available at <https://lccn.loc.gov/2017008317>

ISBN 978-1-107-11681-8 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy
of URLs for external or third-party internet websites referred to in this publication
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Contents

<i>List of Contributors</i>	page vii
<i>Acknowledgments</i>	viii
<i>Note on Text Quotations</i>	ix
Introduction	I
<i>Chris L. Firestone, Nathan A. Jacobs, and James H. Joiner</i>	
PART I KANT AND GOD	11
1 Practical Cognition of God	13
<i>James J. DiCenso</i>	
2 The Birth of God and the Problem of History	35
<i>Pablo Muchnik</i>	
3 The Kantian <i>Summum Bonum</i> and the Requirements of Reason	57
<i>James H. Joiner</i>	
4 Kant and the Experience of God	79
<i>David Bradshaw</i>	
PART II KANT AND RELIGION	97
5 Religious Assent and the Question of Theology: Making Room for Historical Faith	99
<i>Lawrence Pasternack</i>	
6 Kant versus Christianity?	119
<i>Leslie Stevenson</i>	
7 Divine Agency and Divine Action in Immanuel Kant	138
<i>William J. Abraham</i>	

8 Kant and the Problem of Divine Revelation: An Assessment and Reply in Light of the Eastern Church Fathers <i>Nathan A. Jacobs</i>	159
PART III KANT AND REDEMPTION	181
9 What Perfection Demands: An Irenaeian Account of Kant on Radical Evil <i>Jacqueline Mariña</i>	183
10 Atonement and Grace in Kant: Some Reflections <i>Keith Yandell</i>	201
11 Christology . . . within the Limits of Reason Alone?: Kant on Fittingness for Atonement <i>Thomas H. McCall</i>	213
12 Rational Religious Faith in a Bodily Resurrection <i>Chris L. Firestone</i>	228
<i>Bibliography</i>	249
<i>Index</i>	255

Contributors

WILLIAM J. ABRAHAM is Albert Cook Outler Professor of Wesley Studies at SMU Perkins School of Theology

DAVID BRADSHAW is Professor of Philosophy at the University of Kentucky

JAMES J. DICENSO is Professor in the Department for the Study of Religion at the University of Toronto

CHRIS L. FIRESTONE is Professor of Philosophy at Trinity International University

NATHAN A. JACOBS is Visiting Scholar of Philosophy at the University of Kentucky

JAMES H. JOINER is Lecturer in Philosophy at Northern Arizona University

JACQUELINE MARIÑA is Professor of Philosophy at Purdue University

THOMAS H. MCCALL is Professor of Biblical and Systematic Theology at Trinity Evangelical Divinity School

PABLO MUCHNIK is Associate Professor of Philosophy at Emerson College

LAWRENCE PASTERNAK is Professor of Philosophy at Oklahoma State University

LESLIE STEVENSON is Honorary Reader in Logic & Metaphysics at the University of St Andrews

KEITH YANDELL is former Julius R. Weinberg Professor of Philosophy at the University of Wisconsin