

## **Kant's Questions: What is the Human Being?**

Patrick R. Frierson

June, 2009

The present version of this book is a draft. Please do not quote without permission.

Comments are greatly appreciated and may be sent to [frierspr@whitman.edu](mailto:frierspr@whitman.edu)

## Chapter Four: Human Evil and Human History

In the last chapter, we saw that Kant has a detailed empirical account of human beings. While this account does not rise to the level of a “science” in Kant's strict sense, it qualifies as a highly systematic account of universal human characteristics. In this chapter, we will look at two further and related aspects of Kant's empirical account of human beings. These aspects flesh out Kant's empirical account of the human being, and also complete the unfinished business left by the *Critique of Judgment* with respect to the question of what we may hope for humanity as a species.<sup>1</sup> First, we will look at Kant's account of human evil. Kant argues that that human beings are evil “by nature” and also that human evil is “radical” in that it affects the very root of all human choices. Despite this apparently glum assessment of humanity, Kant endorses a realistic hope for human goodness. Second, we will look at one component of this hope, Kant's philosophy of human history, beginning with the emergence of human beings as a new kind of animal with a rational nature and a moral vocation and progressing towards a future of perpetual peace amongst nations and increasingly cosmopolitan political, ethical, and social life for the human race.

### *I. Radical Evil in Human Nature*

#### *a) “The Human Being is Evil by Nature”*

Kant discusses human evil in his *Anthropology* (7:324f.) and in various lectures and notes on ethics, anthropology, and religion, but his most sustained discussion of it takes place in *Religion within the Boundaries of Mere Reason*, a work in which Kant aims “to make apparent the relation of religion to a human nature partly laden with good dispositions and partly with evil ones” (6:11). The book starts with an argument for the existence of human evil, but this argument is complicated by what appear to be contradictory claims. At times, Kant seems to rule out the possibility of knowing anything about one's moral status at all, saying such things as that “we can never, even by the strictest examination, completely plumb the depths of the secret incentives of our actions” (4:407) and that

A human being's inner experience of himself does not allow him so to fathom the depths of his heart as to be able to attain, through self-observation, an entirely reliable cognition of the basis of the maxims which he professes.<sup>1</sup>

Elsewhere, Kant suggests that even if we know that human beings are evil, we should avoid drawing attention to this evil, especially in others: “It is a duty . . . not to take malicious pleasure in exposing the faults of others . . . but rather to throw the veil of philanthropy over their faults, not merely by softening our judgments but by keeping these judgments to ourselves” (6:466). When Kant does propose arguing for human evil, he claims that it “can only be proved [by] anthropological research” (6:25) and that “the existence of this propensity to evil in human nature can be established through experiential demonstrations” (6:35, see too 6:32), but he also insists that “the judgment that an agent is an evil human being cannot reliably be based on experience” (6:20). And even insofar as he does appeal to experience, Kant sometimes seems to argue directly from “the multitude of woeful examples that the experience of human *deeds* parades before us” (6:32-33), but elsewhere insists that

---

<sup>1</sup> Cf. 11:429 where Kant makes explicit that his *Religion* completes the answer to this question.

his claim that “the whole species” is evil can be justified only “if it transpires from anthropological research that the grounds that justify us in attributing . . . [evil] to human beings . . . are of such a nature that there is no cause for exempting anyone from it” (6:25), which suggests that mere examples of evil, even if widespread, are not sufficient. From a quick look at these passages, it becomes unclear whether there can even be an argument for human evil, and if there is, whether that argument is a priori or empirical.<sup>2</sup>

Fortunately, things are not as hopeless as they seem, and Kant's various statements in fact fit together to endorse a complicated but plausible anthropological defense of human evil. The key to putting together Kant's argument comes at the very beginning of his *Religion*, in “Part One: Concerning the Indwelling of the Evil Principle alongside the Good, or, Of the Radical Evil in Human Nature.” There Kant claims,

We call a human being evil . . . not because he performs actions that are evil . . . , but because these are so constituted that they allow the inference of evil maxims in him . . . . In order, then, to call a human being evil, it must be possible to infer a priori from a number of consciously evil actions, or even from a single one, an underlying evil maxim. (6:20)

Kant's argument for evil involves both an empirical component (the experience of “a number of evil actions”) and an a priori component that justifies the inference from these to the “evil maxim” that underlies them. The rest of this section unpacks this argument and its implications for the nature of human evil.

The passage above implies that one *can* infer one's maxims from one's actions. While this might seem to contradict the claims quoted above about the impossibility of self-knowledge, Kant is actually remarkably consistent. One's moral status is inscrutable because any action that accords with the moral law might also be done from some (hidden) inclination and because human beings have a self-deceptive tendency to overestimate the moral worth of our actions.<sup>3</sup> But whenever Kant emphasizes the inscrutability of humans' motives, he rightly emphasizes only that we can never know that our maxims are *good*.<sup>4</sup> But with moral evil, the case is different. While there are no actions that cannot be done from bad motives, there are some actions that cannot be done from good motives. Kant's reference, in the above quotation, to “*actions* that are evil” and his specification of these as “contrary to law,” is important. Generally, for Kant, it is maxims rather than actions that are good or evil. But there are “*actions* . . . contrary to duty” (4:397), and in his *Metaphysics of Morals*, Kant articulates a political theory based on the intrinsic wrongness of actions that cannot “coexist with everyone's freedom in accordance with a universal law” (6:231). Precisely because these actions are wrong regardless of the end for which one performs

---

<sup>2</sup> For various views on Kant's account of and arguments for radical evil, see Allison 1990, 2001 (review of Woodxxx); Frierson 2003; Grenberg 2006; Wood 2009 and 2000:287

<sup>3</sup> There is another, more metaphysical, reason for inscrutability. Since the maxims that are relevant to assessing moral worth are freely chosen and thus “noumenal,” they can never be objects of “knowledge” in Kant's strict sense. For discussion of this point, see Frierson 2003.

<sup>4</sup> In the passages quoted above, when Kant says that “we can never . . . plumb the depths of the secret incentives of our actions” (4:407), this is a response to the fact that “We like to flatter ourselves with the false claim to a more noble motive” (4:407). And when Kant insists that a human being cannot “fathom the depths of his heart,” (R 6:63), the problem here is that we cannot tell whether our maxims have the “purity and stability” necessary to be morally good.

them, one can legitimately infer bad underlying maxims from the performance of such actions.

Moreover, because the inscrutability of one's moral status comes in part from the tendency to self-flattery, it is implausible that one would pretend to a motive *less* noble than one's actual motive, so when one finds in self-scrutiny that one's motive for an action is evil (what Kant calls "consciously evil"), one can reasonably trust that there is no underlying righteous motive for that action, even if the action is not strictly "contrary to law." The inscrutability of human motivation is thus asymmetrical: one can never know that a person – including oneself – is morally *good*, but one can know that people are morally evil.

Even if Kant's claims about the inscrutability of ultimate intentions do not preclude knowledge of human evil, though, how can Kant make inferences from experience to the existence of human evil given that "the judgment that an agent is an evil human being cannot reliably be based on experience" (6:20)? Here it is important to see that neither experience nor a priori arguments alone are *sufficient* for Kant's proof of evil (hence both are necessary). Experience of actions contrary to duty would not be sufficient for ascribing an evil will to human beings without an argument that links those actions to evil maxims. So what is needed in order to move from evil actions to evil maxims? Given evil actions (either because they are wrong in themselves or because they are "consciously evil"), one knows that *if* those actions are grounded in freely chosen maxims, then the maxims are evil. So in order to establish a connection between evil actions and evil maxims, all that is needed is an argument that human beings are free agents who choose in accordance with maxims that can ground evil actions such as those found in experience. Kant has already provided much of this account in his transcendental anthropology. In his *Religion*, Kant insists that "experience can never expose the root of evil in the supreme maxim of a free power of choice, for, as intelligible deed, the maxim precedes all experience" (6:39n), but Kant's transcendental anthropology has already shown, a priori, that human actions are the phenomenal expressions of noumenal, free choices; we act only on the basis of incentives that we freely incorporate into maxims of action.<sup>5</sup>

In *Religion*, Kant adds to his general account of human transcendental freedom an account of the specific structure of the fundamental maxim that grounds evil actions. In particular, Kant makes two important additions to the account of free choice found elsewhere in his Critical philosophy. First, he argues that human choices must be grounded in a basic maxim that is either fundamentally good or fundamentally evil; no middle ground is possible. As Kant puts it,

[I]f [someone] is good in one part [of life], he has incorporated the moral law into his maxim. And were he . . . to be evil in some other part, since the moral law of compliance with duty in general is a single one and universal, the maxim relating to it would be universal yet particular at the same time: which is contradictory. (R 6:24-5)

Because morality requires *unconditional* and *universal* compliance (4:416), one who sometimes but not always acts in conformity with the moral law never *really* makes the

---

<sup>5</sup> We discussed this principle in terms of the "Incorporation Thesis" in chapter two. For a discussion of the relationship between empirical claims about human motivation and the ascription of transcendental freedom to them, see Frierson xxx (Empirical Markers).

moral law his ultimate motive, since any law whose application depends upon circumstances cannot be the moral law.<sup>6</sup>

Second, Kant connects the account of humans' free finitude from his transcendental anthropology with the account of human predispositions in his empirical anthropology. As we saw in the last chapter, Kant's conception of a "predisposition" has wide application, covering all basic human powers and the instincts and propensities that direct the faculty of desire. In *Religion*, Kant employs this notion of a predisposition in his discussion of a fundamental "predisposition to the good" in human nature that consists of three distinct "elements of the determination of the human being," animality, humanity, and personality (6:26-7, cf. 7:322-5). The predisposition to animality includes basic instincts and even propensities to various inclinations that support human life. In particular, Kant emphasizes instincts for self-preservation, for propagation of the species (sexual instinct), and "for community with other humans, i.e. the social drive" (6:26). The predisposition to humanity includes our innate tendency to compare ourselves with others and thereby a propensity to "the inclination to *gain worth in the opinion of others*" (6:27). Finally, the predisposition to personality is "susceptibility to respect for the moral law *as of itself a sufficient incentive in the power of choice*" (6:27). As we saw in the last chapter, higher cognitions are generally capable of determining the power of choice, but there is a particular predisposition by virtue of which a *pure* higher cognition is capable of determining choice.

By subsuming the variety of human volitional predispositions under the general category of "predisposition to the good," Kant emphasizes that no natural human instincts or inclinations are evil in themselves: "the ground of evil cannot . . . be placed . . . in the sensuous nature of the human being" (6:34). But because the good predispositions of human beings include some that are not *unconditionally* or *morally* good, there is a basis in human nature for evil. Kant goes on to unpack that account.

The human being (even the worst) does not repudiate the moral law . . . The law rather imposes itself on him irresistibly, because of his moral predisposition; and if no other incentive were at work in him, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice . . . He is, however, also dependent upon the incentives of his sensuous nature because of his equally innocent natural predisposition, and he incorporates them too into his maxim . . . Hence the difference, whether the human being is good or evil, must not lie in the difference between the incentives that he incorporates into his maxim . . . but in their *subordination* . . . : *which of the two he makes the condition of the other*. It follows that the human being (even the best) is evil only because he reverses the moral order of his incentives in incorporating them into his maxims. (6:36, see too 6:32)

In this important passage, Kant lays out the essence of his account of human evil. Importantly, the account can be read both in terms of transcendental freedom and in terms of empirical anthropology. The transcendental reading is crucial since in the absence of a

---

<sup>6</sup> It may be worth noting an important difference here between the moral law and other motivational principles. Inclinations and instincts are often operative only in particular contexts, and principles based on inclination/instinct can be even more mitigated. For instance, one who makes the satisfaction of sexual desire their overriding end may not feel sexual desire in particular contexts, and may even refrain from acting on a present desire in order to satisfy more sexual desire later. While nonmoral incentives are thus often highly context dependent, the moral law is essentially universal.

transcendental perspective on the subordination of moral to nonmoral incentives, no empirical claim can imply anything about human *evil*. For Kant, “In freedom alone is evil.”<sup>7</sup> From the perspective of freedom, when one looks at one’s action from-within, what Kant claims here is that in all choices, we have concern both for morality and for well-being (defined in terms of animal and social inclinations), but that ultimately, we subordinate one concern to the other. Our free (noumenal) nature is constituted by whether we unconditionally prioritize the moral law to nonmoral concerns or whether we allow nonmoral concerns of sufficient weight to trump the moral law. This aspect of Kant’s account depends crucially upon the a priori account of morality from Kant’s transcendental anthropology, within which Kant shows both that human beings are transcendently free and that morality requires unconditional obedience. Here, Kant uses these claims to argue that because morality requires *unconditional* obedience from a transcendently free will, *any* subordination of moral to nonmoral concerns is *wholly* evil.<sup>8</sup>

But Kant’s argument for human evil is not merely an argument directed towards helping readers recognize evil from-within. He also aims to make an empirical-anthropological point, that human beings are evil *by nature*. From the standpoint of empirical anthropology, the key passage above can be read as a step towards completing his empirical account of human nature. Human beings have various predispositions that can be classified in terms of animal instincts, social inclinations for recognition, and moral interests. But a complete empirical account of human beings must discern how these needs interact in cases when more than one is active. Kant thus looks for empirical evidence to suggest that the empirical character of human volition is structured such that moral grounds are inactive when they conflict with sufficiently strong nonmoral grounds. What is this empirical evidence? At the most fundamental level, Kant finds in the “multitude of woeful examples” of human misdeeds evidence for the empirical claim that humans’ volitional structure is such that moral incentives are subordinated to nonmoral ones. To show that this corrupt volitional structure cannot be ascribed to societal influence alone, Kant points out the presence of “unprovoked cruelty” in the state of nature, and to show that this corrupt volition is not limited to uncivilized savages, he lays out a host of vices of the “civilized state,” including “secret falsity even in the most intimate friendship . . . , a propensity to hate him to whom we are indebted . . . , many other vices yet hidden under the appearance of virtue, [and] those of which no secret is made [wherein] . . . someone already counts as good when *his evil is common to a class*” (6:33). This evidence gives Kant grounds for the empirical-anthropological claim that humans’ particular choices are grounded in a power of choice within which moral predispositions are subordinated to nonmoral ones. Humans have both predispositions that make evil possible and a volitional *structure* within which the moral predisposition is made inactive by sufficiently strong sensuous incentives such that given our

---

<sup>7</sup> *Reflexion* 5541, 18:212; cf. *Moralphilosophie Collins*, 27:295.

<sup>8</sup> Kant often even seems to present his argument for evil as primarily an argument addressed to individual, free agents. He emphasizes that we “notice (at least within ourselves) [actions that] are consciously contrary to [moral] law” (6:20) and gives examples of vices, especially those in civilized society, that his readers will – reluctantly – have to admit as applying to themselves. When he concludes his argument for evil in human nature, he even emphasizes this personal aspect of it. Whether or not “every man has his price, for which he sells himself” is something, Kant suggests, that “everyone can decide for himself” (6:39).

transcendental freedom (established by Kant's transcendental anthropology), human beings are shown to be evil. Transcendentally speaking, there is no *necessity* for human beings to have this volitional structure; it is contingent upon transcendently free choice. But empirical speaking, when one seeks to discern human nature based on empirical evidence, there is good reason to think that human volition subordinates *pure* higher volition to *impure* higher volition. And given that Kant's transcendental anthropology shows this empirical character to be grounded in free choice, there is reason to describe this subordination as "evil."

In the end, then, Kant's argument for evil in human nature is simple in outline and rich in detail.

1. In widely varying circumstances, human beings perform actions that contradict the moral law and/or consciously perform actions that are immoral.
2. Human actions result from the influence of empirical causes through ordered predispositions that determine how empirical causes effect particular actions.
3. Human beings have both a moral predisposition according to which they can be motivated by the moral law and nonmoral predispositions to pursue natural and social goods.
4. The moral law is essentially unconditional, requiring stable and pure adherence.
5. Thus, human behavior is characterized by a prioritization of nonmoral predispositions over the moral predisposition.
6. Humans' empirical behavior and character express their transcendently free choices.
7. Thus, human beings are morally evil.

The first three premises are empirical generalizations, of different levels of complexity. The first is a straightforward generalization of observations about human beings. The second and third<sup>9</sup> generalize an anthropological explanatory model from a number of cases in which it has been observed to be a good explanation for observed human behavior. These premises are developed in much greater empirical detail, as we showed in chapter three. The fourth premise is a moral premise, a part of Kant's "transcendental" anthropology of the faculty of desire. The evidence for this claim is thus a priori. In the context of a properly empirical argument, this a priori premise would have to be taken as stipulative, merely a claim about what "moral" means, in which case it would add no a priori content to the argument of the first three premises, but only specify what premise (3) means. In that sense, the preliminary conclusion at (5) could be taken to be a properly empirical-anthropological conclusion. Given that prioritizing the moral predisposition would involve (by definition) consistency in following that predisposition, it is clear from premises (1)-(3) that human beings act according to a complex structure of predispositions within which the "moral" predisposition is subordinated to others. And in that sense, (5) expresses an empirical fact about human nature. But premise (6) is essentially transcendental; there is no empirical

---

<sup>9</sup> In his *Critique of Practical Reason*, Kant also offers an "a priori" argument for his third premise (see 5:78). Given our obligations, we can know a priori that we have a capacity to act out of respect for the moral law, which implies that we have a *power* to do so. For the purposes of his proof in the *Religion*, however, this premise is taken as part of Kant's empirical anthropology.

evidence for humans' status as free grounds of their empirical characters. Given this premise, however, the prioritization of nonmoral predispositions over the moral predisposition that was shown to be a part of human nature is also revealed as an expression of truly moral evil. The conclusion which is both transcendental (a priori) and empirical, is that human beings are evil by nature, that is, that moral evil can be ascribed to every member of the human species.

Having shown that human beings are evil, Kant then elaborates on the nature of that evil. Most importantly, Kant emphasizes that human evil is "radical" in that "it corrupts . . . the subjective supreme ground of all maxims" (6:37). By this Kant means that the "maxim" by which humans subordinate moral to nonmoral incentives is their most fundamental maxim. In general, humans act in accordance with various principles (maxims) of action, but these can be ordered in a kind of hierarchy. To take one of Kant's own examples, one might act on the maxim "when I believe myself in need of money I shall borrow money and promise to repay it, even though I know that this will never happen" (4:422), but this maxim is merely a particular application of more general maxims such as "I will trust my own assessments of my needs" and "when I can make use of others to satisfy my needs, I will do so (regardless of whether I treat them as ends when I do so)," and this latter maxim is a more specific application of an even more general maxim that Kant explains in terms of the relative subordination of inclinations and morality, something like "I will obey the moral law only insofar as doing so is compatible with satisfying other desires, and I will always seek to satisfy some nonmoral desires." This maxim, Kant suggests, is the fundamental guiding maxim of an evil human being's life, and all other maxims are merely more specific applications of this one to particular cases where one's inclinations and/or the moral law are in play. Because this corrupt maxim lies at the root of all one's choices, Kant refers to human beings as "*radically* evil."

In laying out this account of radical evil, Kant also clarifies some important details about the nature of evil. For one thing, he argues that radical evil is not only "itself morally evil, since it must ultimately be sought in a free power of choice" (6:37), but also that it is tied to a "natural propensity to evil" that structures the particular evil choices that human beings make.<sup>10</sup> Moreover, the source of radical evil in *choice* implies that radical evil "cannot be placed, as is commonly done, in the sensuous nature of the human being and in the natural inclinations originating from it" (6:34-5). At one level, this claim should be quite obvious. As a moral category, evil cannot properly be located in the human being *qua* object of empirical investigation but must be traced to the free, noumenal agent that grounds empirically observable behavior. But Kant's claim here also shifts the focus even within empirical character. The empirical expression of radical evil is not in the lower faculties – the senses and inclinations – but in the higher faculties, especially in the higher faculty of desire. Human agents, even as empirically observed, have a capacity – what Kant calls "the power of choice" – to act from principles, and it is the way in which this capacity is used that gives empirical evidence of freely chosen evil.

---

<sup>10</sup> Many commentators see this propensity to evil as a *precondition* of radical evil (cf. Allison 1990, Wood 2000), but I see Kant as portraying the propensity to evil as both a *consequence* of humans' radical evil and as a *ground* of further, particular evil choices (see Frierson 2003).

Kant also describes three ways in which evil might express itself in one's choices: frailty, impurity, and depravity. The first involves merely a lack of character, or what Kant calls in his anthropology lectures a "bad character," an "inability to act according to principle" (25:650). In these cases, the principles of one's higher faculty of desire are good, but when it comes to actually acting, these principles do not actually determine one's actions. As we noted in the last chapter, there can be conflicting underlying grounds of action, and often one or more powers are "dead" or "inactive" while others are active in effecting a transition to a new mental state or an action. Those with frail wills understand the principles according to which they should act, and the character of their higher faculty of desire is such that "I incorporate the good (law) into the maxim of my power of choice, but this good . . . is subjectively the weaker (in comparison with inclination) whenever the maxim is to be followed" (6:29). In the paradigm cases of frailty, one's higher faculty of desire is properly oriented such that, if active, it would cause one to do what is right. But when the relevant moment comes, the higher faculty of desire is weaker than inclination (the lower faculty of desire) and hence inactive. We might imagine more complex cases, where inclination does not wholly overcome the higher faculty of desire but prevents it from actually bringing about effects in accordance with its character; here the inclination subverts the normal operation of the higher faculty, rather than preventing its operation altogether. In these cases, the propensity to be governed by fixed principles is not fully developed. And the account of evil in *Religion* argues that this empirically observable badness of character can be ascribed to a free (noumenal) choice to subordinate morality to inclination.

The other two forms of evil are more straightforward since they both involve acting in accordance with principles of a corrupted higher faculty of desire. "Impurity" occurs when one's "maxim is good with respect to its object . . . [but] has not . . . adopted the law *alone* as its *sufficient* incentive" (6:30). One who is impure generally chooses what is morally required, but always only because it is *both* morally required *and* conducive to satisfying other desires. This *conditional* adherence to the moral law is, as Kant's argument for rigorism shows, no adherence at all. The final form of radical evil is "depravity," which involves a specific "propensity of the power of choice to maxims that subordinate the incentives of the moral law to others (not moral ones)" (6:30). One who is depraved might still often act in ways that seem moral, but the depraved person's power of choice is structured by a fundamental commitment to nonmoral desires, regardless of whether these are morally permitted or not.

Importantly, Kant rejects the possibility of what he calls "diabolical" evil, the "disposition . . . to incorporate evil *qua* evil . . . into one's maxim" (6:37). For Kant, even the most evil person is not motivated by evil *as such*. There are no *immoral* desires in human beings, since all desires can be traced to one form of the predisposition *to the good*. One is evil when one allows *nonmoral* desires to trump the moral law, not when one chooses evil as such. Thus Kant does not allow the possibility of cases like St. Augustine's famous theft of pears "not to eat for ourselves, but simply to throw to the pigs[, for] our real pleasure consisted in doing something that was forbidden."<sup>11</sup> For Kant, human beings do not have a desire to do what is morally forbidden *per se*. Evil arises only from putting *nonmoral* desires ahead of our innate moral predisposition.

---

<sup>11</sup> Xxx § 4, p. 47 in Penguin Classics Edition.

Finally, it is important to note that in all of these cases, one's radical evil need not imply that one *always* chooses contrary to the moral law. To be evil is to be disposed to allow the moral law to be overridden *given a sufficient sensuous incentive*. At the end of a section explaining that "the human being is by nature evil," Kant adds,

A member of the English Parliament exclaimed in the heat of debate: "Every man has his price, for which he sells himself." If this is true (and everyone can decide for himself), if nowhere is a virtue which no level of temptation can overthrow, if whether good or evil wins us over only depends on which bids the most and affords the promptest pay-off, then what the Apostle says might indeed hold true of human beings universally, "There is no distinction here, they are all under sin – there is none righteous (in the spirit of the law), no, not one." (6:39)

Frailty, impurity, and even depravity all involve, in different ways, a subordination of the moral law to nonmoral desires. But one can be radically evil and still often do what is good, if one does what is good only because the price of doing good is, in a particular case, not too high.

### ***b) The Problem of Radical Evil***

Kant's claim that human beings are radically evil raises a serious problem at the intersection of transcendental and empirical anthropology, a problem that Kant spends the rest of his *Religion* trying to solve. Put simply, the problem is that evil seems inextirpable.

This evil is *radical*, since it corrupts the grounds of all maxims; as a natural propensity, it is also not to be *extirpated* through human forces, for this could happen only through good maxims – something that cannot take place if the subjective supreme ground of all maxims is presupposed to be corrupted. (6:37, cf. 6:45)

Because evil lies at the root of human choice, one "can" not extirpate it through that same (evil) power of choice. (One might think that being evil "by nature" precludes transcendental freedom and thus moral responsibility, but for Kant, we are evil "by nature" *through* our free choice. Radical evil is a *consequence* of humans' use of their transcendental freedom; it is only because we *freely choose* to subordinate moral to nonmoral incentives that such a choice can be considered evil. But given that we freely choose evil *as the basis of all of our other choices*, it seems impossible for that freedom to be used to rid us of radical evil.)<sup>12</sup> That is, it seems impossible that we can both choose evil and choose to extirpate that evil. And the problem of radical evil is made even worse by our evil tendency to cultivate a propensity to evil. Human beings not only choose in evil ways but also deliberately cultivate both themselves and their environment (especially their social environment) in order to promote the easy exercise of evil tendencies. One whose evil manifests itself primarily in a frail will, for example, has cultivated a weakness of character that will be hard to overcome even if one tries to do so. And one who is depraved has developed patterns of self-deceptive moral justification and corrupting social interaction that will make it difficult for good intentions to fully overcome selfishness. Finally, the problem is even more acute because no matter how good one might be able to become, one *has chosen* in accordance with the evil

---

<sup>12</sup> The case here is similar to the case of "passions," of which Kant says in his *Anthropology* that "they are incurable because the sick person does not want to be cured and flees from the dominion of principles by which alone a cure could occur" (7:266).

principle, so one can *never* be a person who *always* chooses in accordance with the moral law:

however steadfastly a human being may have persevered in such a [good] disposition in a life conduct conformable to it, *he nevertheless started from evil*, and this is a debt which is impossible for him to wipe out. (6:72)<sup>13</sup>

Altogether, not only is one's choice oriented in such a way that one rejects moral reform (*radical evil*), but even if one were somehow to begin such a process of reform, one would have to contend with self-wrought influences that make morally upright action difficult (the *propensity* to evil), and even if one somehow overcame these influences, one would never have a life that was *wholly good from start to finish* (one *started from evil*).

The inextirpability of radical evil *need* not pose a philosophical problem for Kant. Kant might just say that the empirical evidence shows that human beings are radically evil through their own fault, and so much the worse for us. Even Kant's claim that "ought implies can" need not be compromised by the claim of radical evil because radical evil is an evil that is self-wrought through our own freedom. Humans *could* be good, but we universally (but not necessarily) choose not to be; evil's "inextirpability" is *due to* choice, not a constraint on it. But Kant is committed to a moral vision that goes beyond mere insistence upon moral responsibility for wrongdoing: "In spite of the fall, the command that we ought to become better human beings still resounds unabated in our souls; consequently, we must also be capable of it" (6:45). Even if radical evil does not eliminate moral responsibility, it does seem to deny the real possibility of moral reform, which would undermine the force of the obligation to *improve*. In response, Kant defends *moral hope*, the possibility of reforming oneself morally despite one's radical evil. And this commitment to hope generates a problem: how can one reconcile moral rigorism, radical evil, and moral hope?<sup>14</sup>

At one important level, Kant does not even try to explain how moral reform is possible given radical evil. When he points out that evil cannot be extirpated, he adds "through human forces" (R 6:37) and then suggests, "Some supernatural cooperation is also needed to [a person] becoming good or better" (R 6:44).<sup>15</sup> But this "supernatural cooperation" is ultimately beyond rational comprehension and even practical use.<sup>16</sup> The main role of such cooperation, which Kant calls "grace," is simply to reinforce the need for human beings to do their part to "make themselves antecedently worthy of receiving it" (R 6:44, cf. 6:118). Kant emphasizes that the inscrutability of grace is no greater than the inscrutability of

---

<sup>13</sup> There are obvious connections between Kant's account of evil and the Christian doctrine of original sin. Although I do not read the *Religion* as merely a rational reconstruction of Christian doctrine, the similarities are undeniable. For more on this theme, see Quinn 1984, 1986, and 1990; Adams 1998; and Hare 1996.

<sup>14</sup> For discussion of this problem in similar terms, see Hare 1996.

<sup>15</sup> For discussion and further references, see Frierson, 2003, pp. 114-22.

<sup>16</sup> E.g. R 6:117-8, 191; SF 7:43-4. Admittedly, Kant is conflicted about the practical use of this supernatural cooperation. On the one hand, the problem of radical evil seems to preclude human beings from choosing well antecedent to some sort of grace. But on the other hand, Kant insists that one "must accept this help (which is no small matter), i.e. he must incorporate this positive increase of force into his maxim: in this way alone is it possible that the good be imputed to him, and that he be acknowledged as a good human being" (6:44). For discussion, see Marina xxx, Frierson 2003, and Adams xxx.

freedom<sup>17</sup> and even that humans' continuing recognition of their moral obligations shows an enduring "germ of goodness . . . that cannot be extirpated or corrupted" (R 6:45-6; see too 6:49; SF 7:43, 58-9; A 7:322). The enduring germ of goodness shows that all people still have a capacity for goodness, and one's freedom gives an enduring but inexplicable hope that this capacity can still be used well. Of course, none of these claims about inscrutability actually address the central problem of radical evil, and even Kant's appeal to grace, absent further explanation and especially given his insistence upon humans' need to be *antecedently* worthy of that grace, does not show *how* grace solves the problem of evil.

But Kant's theoretically inadequate discussion of radical evil does highlight the proper stance to take towards the problem. Given his transcendental anthropology of cognition, Kant is surely correct that the metaphysical mechanisms by virtue of which radical evil might be overcome will never be understood by human beings. But the problem of radical evil is not, fundamentally, a metaphysical problem but a practical one. What ought one *do* in light of radical evil and what may one *hope* with respect to it? In one sense, the practical aspect of radical evil is easy to address. If evil is a free choice to subordinate the moral law to nonmoral desires, then what one must *do* is subordinate one's nonmoral desires to the moral law. Here Kant can do little more than exhort people to goodness and warn against self-deceptive despair or weakening of moral demands.<sup>18</sup> But radical evil is also a self-wrought *tendency* to act immorally, and it is, moreover, a tendency that is in evidence *in the human by nature*. And these aspects of radical evil require some account of the grounds for moral hope in the human species as a whole as well as an account of how one can work to undo and arm oneself against one's self-wrought evil tendencies. Kant deals with the former task in his sophisticated philosophy of human history, a history situated in the context of radical evil but one that justifies hope in humanity's future. Kant deals with the second task in his "moral anthropology," which deals with "the subjective conditions in human nature that hinder people or help them in *fulfilling* [moral] laws . . . , with the development, spreading, and strengthening of moral principles" (6:217). The rest of this chapter focuses on Kant's philosophy of history. Kant's moral anthropology will be discussed in chapter six.

## ***II. Human Beings as an Historical Species***

Kant is not generally known for his philosophy of history, and a historical conception of human beings can seem to be at odds with other important aspects of Kant's philosophy, especially his concern with individual freedom. Nonetheless, during the height of work on the transcendental philosophy that made him famous, Kant was also writing a series of papers on human history. In the first of these, his "Idea for a Universal History with

---

<sup>17</sup> "Since the fall from good into evil . . . is no more comprehensible than the ascent from evil back to good, then the possibility of this last cannot be disputed" (R 6:45, see too 6:191).

<sup>18</sup> To be fair, Kant's doctrine of grace can do a bit more than mere exhortation. Kant offers a detailed account of justificatory grace whereby suffering in one's life atones for the misdeeds of one's past and within which progress towards the good comes to be counted as perfect goodness. For discussion of these more specifically religious aspects of Kant's justification of moral hope, see Quinn xxx, Marina, xxx, Michaelson xxx, Frierson 2003, Frierson forthcoming, xxx ("Kant's Religion"). My focus in this book will be on more specifically "anthropological" answers the question of what one may hope, both through Kant's conception of historical progress and through his account of the moral anthropology that radical evil makes necessary and grace makes possible.

a Cosmopolitan Aim,” Kant claims that human beings have “predispositions . . . [that] develop completely only in the species, but not in the individual” (8:18). Kant’s empirical account of human beings is not limited to the nature and development of individual human lives but must include an account of the historical development of the human race as a whole.

Methodologically, Kant’s history starts from the empirical anthropology discussed in the last chapter. Like empirical anthropology generally, the study of human history is primarily empirical. Kant begins his “Idea” essay by emphasizing this point:

Whatever concept one may form of the *freedom of the will* with a metaphysical aim, its *appearances*, the human actions, are determined just as much as every other natural occurrence in accordance with universal laws of nature. History, which concerns itself with the narration of these appearances, however deeply concealed their causes may be, nevertheless allows us to hope from it that if it considers the play of the freedom of the human will *in the large*, it can discover within it a regular course. (8:17)

But history is not “mere empirical groping without a guiding principle of what to search for” (8:161). Kant’s account of predispositions provides this guiding principle. While the empirical anthropology of the previous chapter focused on predispositions as bases of the causal powers by virtue of which various causes give rise to psychological effects, Kant’s philosophy of history studies these predispositions *teleologically*. In his *Critique of Judgment*, Kant argued that organic life could be interpreted in terms of predispositions that serve purposes within an organism: all organisms are “conceived of teleologically under the concept of a natural end” (5:376). For the purposes of his philosophy of history, Kant takes this account of the natural purposiveness of organisms further in two important ways. First, Kant argues that internal natural purposiveness gives rise to a question about the “external purposiveness” of an organism, “that in which one thing in nature serves another as the means to an end” (5:425). Once something is posited as a natural end, “one can . . . ask, why does it exist?” and unless “it is not only an end, but also a *final end*,” one must think of its existence as for the sake of something else. From this general claim, Kant develops an ecological approach to biology, according to which one can see how organized beings function in the context of whole systems of such beings (see 5:426-7), but he also argues that “we have sufficient cause to judge the human being . . . as the ultimate end of nature here on earth” (5:429). Second, in his writings on history in particular, Kant emphasizes an important presupposition of ascribing any purposive structure to an organic being:

*All natural predispositions of a creature are determined sometime to develop themselves completely and purposively. . . . An organ that is not to be used, an arrangement that does not attain to its end, is a contradiction in the teleological doctrine of nature. For if we depart from that principle then we no longer have a lawful nature but a purposelessly playing nature; and desolate chance takes the place of the guideline of reason.* (8:18)

For most animals, this teleological assumption has implications only for the study of individual organisms. To identify a feature of an organism as a physical or behavioral predisposition, one must assume that it serves a purpose, and this implies that at some point in the development of the organism, the feature will develop in the way needed to serve that purpose. For human beings, however, some predispositions are not fully realized in the life of any single person. For example, the full development of human reason in the arts,

sciences, and politics does not occur in any single individual's life but happens over the course of the history of the species. But insofar as one still treats capacities such as reason as natural predispositions, one must apply the same regulative principles to them as to other predispositions; one assumes that they will develop toward their end. And this assumption provides an "Idea" that can underlie a rationally guided but empirically based history that looks for the ways in which humans' natural predispositions unfold over time.<sup>19</sup>

Kant's treatment of the earliest human history is laid out in the quasi-satirical "Conjectures on the Beginning of Human History," which offers a sort of scientific commentary on the story of humans' creation in the book of Genesis. While some philosophers and anthropologists in the 18<sup>th</sup> century sought to show how human beings developed from other primates – the issue of the relationship between the upright posture and human reasoning was a hot topic of the day – Kant starts with "the *existence of the human being . . . in his fully formed state . . . [and] in a couple*" (8:110). By "fully formed" Kant means only that the humans have all of their natural predispositions, not that these are all fully developed, but even this assumption means that Kant does not need to explain, as his student Herder aimed to do, how "*psychology*" arises from "determinate *physiology*," how higher cognitions arise from the contractions and expansions of "irritated little fiber[s]."<sup>20</sup> Instead, Kant starts with primitive rational and sexual beings and shows how humans developed from that stage. In this essay, the key development that inaugurates truly *human* history is "the first development of freedom from its original predisposition in the nature of the human being" (8:109). In his "Idea," Kant had argued that "Nature has willed that the human being should produce everything that goes beyond the mechanical arrangement of his animal existence entirely out of himself" (8:19), and "Conjectures" shows how an animal with the mere *potential* for this sort of free species-development comes to have the actual freedom needed for it.<sup>21</sup>

Kant outlines four steps into the state of true, actualized human freedom. Blending Genesis with Rousseau,<sup>22</sup> Kant first describes how human beings come to desire objects that are not natural objects of instinct:

Instinct . . . must have guided the novice. It allowed him a few things for nourishment but forbade him others . . . As long as the inexperienced human being obeyed this call of nature, he did well for himself. Yet *reason* soon began to stir and sought through comparison of that which gratified with that which was represented to him by another sense than the one to which instinct was bound, such as the sense of sight, as similar to what previously was gratifying, to extend his knowledge of the means of nourishment

---

<sup>19</sup> Xxx: Possibly compress this ¶ to avoid repetition with earlier chapter.

<sup>20</sup> Herder 2002, pp. 196, 189. For further discussion, see xxx

<sup>21</sup> This freedom is still merely empirical, in that it is part of an empirical-historical account of human beings. It is a sort of capacity to act on the basis of principles (character) rather than mere inclinations. Because this empirical sort of freedom is correlated, for Kant, with transcendental freedom (see Frierson xxx), one might see this as the historical origin of transcendental freedom. Strictly speaking, however, Kant's history of humanity, as a part of his empirical anthropology, cannot describe how *transcendental* freedom emerges.

<sup>22</sup> Kant presents his account as a quasi-commentary on Genesis 3, but its substance is Rousseauian (see especially Rousseau's *Discourse on the Origin of Inequality*).

beyond the limits of instinct . . . The occasion for deserting the natural drive might have been only something trivial; yet the success of the first attempt, namely of being conscious of one's reason as a faculty that can extend itself beyond the limits within which all animals are held, was very important and decisive for his way of being. (8:111-12)

The first stage in the development of human freedom comes when humans' cognitive faculties develop to the point at which they are capable of modifying human desires. Whereas the earliest humans pursued objects to which they were instinctually drawn, at some point human beings decide, for instance, to try "a [new] fruit whose outward look, by its similarity with other pleasant fruits . . ., invited him to the attempt" (8:112). Humans' faculties of desire are no longer wholly at the mercy of their lower, sensory faculties of cognition, but become capable of control by the higher faculty of cognition, by conceptual awareness and principles for action.

Following Rousseau (and Genesis), Kant does not see this first step into freedom as wholly beneficial. The ability to generate new desires includes an ability to generate *unhealthy* desires; one can "concoct desires not only *without* a natural drive . . . but even *contrary to it*" (8:111). Moreover, freedom over desires causes a new *problem* of freedom for human beings:

[The first human] discovered in himself a faculty of choosing for himself a way of living and not being bound to a single one, as other animals are. Yet upon the momentary delight that this marked superiority might have awakened in him, anxiety and fright must have followed right away, concerning how he, who still did not know the hidden properties and remote effects of any thing, should deal with this newly discovered faculty. He stood, as it were, on the brink of an abyss; for instead of the single objects of his desire to which instinct had up to now directed him, there opened up an infinity of them, and he did not know how to relate to the choice of them; and from this estate of freedom, once he had tasted it, it was nevertheless wholly impossible for him to turn back again. (8:112)

When one is motivated solely by the lower faculty of desire, one need barely even think about proper means to one's ends. One pursues objects of desire almost without thinking. But once one is capable of generating new desires through reasoning, one must decide which objects are *worth* pursuing among an apparently infinite expanse of possibilities. But one still lacks any proper framework for making such determinations.

While the first stage in human freedom transformed desires in general (especially desires oriented towards personal physical needs), the second stage transforms the most intense and powerful social instinct in human beings: the sexual instinct. Following Rousseau, Kant sees a fundamental difference between the raw desire for sex and the way in which sexuality plays out in human life. In particular, human beings overlay onto their desire for sexual gratification an interest in the beauty and even personality of the sex object. Picking up on the Biblical reference to Adam and Eve covering themselves with fig leaves, Kant envisages reason's rising influence over human desires. By covering themselves,<sup>23</sup>

---

<sup>23</sup> In the account in "Conjectures," Kant sees this artifice of refusal as used by both Adam and Eve. Elsewhere (cf. Anthropology, Observations, Remarks) Kant suggests that the art of refusal is a distinctively feminine art.

Adam and Eve make themselves more desirable, and the sexual instinct gets infused with ideals of beauty and propriety. Here reason “make[s] an inclination more inward and enduring by withdrawing its object from the senses,” which “shows already the consciousness of some dominion of reason over impulse and not merely, as in the first step, a faculty for doing service to those impulses within a lesser or greater extension” (8:113).

The third step involves the “deliberate *expectation of the future*” (8:113), which requires still higher and more organized interactions between reason and desire. Like the first steps, the effects of this are ambivalent: it “is the most decisive mark of the human advantage of preparing himself to pursue distant ends in accordance with his vocation – but also simultaneously it is the most inexhaustible source of cares and worried which the uncertain future incites and from which all [other] animals are exempt” (8:113).<sup>24</sup>

Finally, in the last stage the human being “comprehended (however obscurely) that he was the genuine *end of nature*” (8:114). Human beings come to see the products of nature as possible instruments for their own use (cf. Genesis 3:21), but they also recognize – albeit obscurely – that every other human beings is “equal participant in the gifts of nature” and thus can rightly make “the claim of *being himself an end*, of also being esteemed as such . . . , and of being used by no one merely as a means to other ends” (8:114). Kant does not, of course, think that the earliest human beings had a worked out theory of human rights, nor even that they actually treated all other human beings as equals. In fact, Kant is well aware that human beings seek to dominate one another and treat one another as mere instruments to one’s personal ends. But this domination among human beings has, according to Kant, a fundamentally different character than the struggle with the rest of nature. Among beings who are all capable of forming plans for themselves on the basis of “a faculty of choosing for himself a way of living” (8:112), influence takes a form either of blameworthy domination or of mutual cooperation.

This emergence into freedom marks only the beginning of Kant’s historical account of human beings. Before emerging into freedom, human beings were distinguished from animals only by latent predispositions to higher cognitive and volitional faculties. But upon becoming truly free, humans could become a truly historical species. At this stage, the importance of the claim that “human . . . predispositions . . . develop completely only in the species” (8:18) comes to the fore, and Kant adds to this a further claim central to his account of human history:

*Nature has willed that the human being should produce everything that goes beyond the mechanical arrangement of his animal existence entirely out of himself, and participate in no other happiness or perfection than that which he has procured for himself free from instinct through his own reason. (8:19)*

A human’s faculty of choosing for himself a way of living combined with his decisive ability to “prepar[e] himself to pursue distant ends in accordance with his vocation” generates the structure of human history, according to which all development of human predispositions occurs by means of humans’ own deliberate work. But Kant almost immediately adds an important caveat to this emphasis on freedom. While human history progresses by means of

---

<sup>24</sup> Again, see Rousseau’s *Discourse on the Origin of Inequality*.

human choices, Nature<sup>25</sup> uses human choices to achieve ends that diverge from the immediate ends of the choices themselves. In particular,

*The means nature employs in order to bring about the development of all [humans'] natural predispositions is their antagonism in society, insofar as the latter is in the end the cause of their lawful order. Here I understand by 'antagonism' the unsocial sociability of human beings, i.e., their propensity to enter into society, which, however, is combined with a thoroughgoing resistance that constantly threatens to break up this society. (8:20)*

Human beings have both a natural inclination to enter into society with others and a desire to exert superiority over others. Elsewhere, Kant explains that humans have a set of “predispositions to *humanity*” that “can be brought under the general title of a self-love that is physical and yet *involves comparison* . . . that is, only in comparison with others does one judge oneself happy” (6:27). Thus one must be in the company of others (in order to compare oneself to them), but because “the inclination to gain . . . equal worth” fast becomes “an unjust desire to acquire superiority for oneself over others” (6:27), there arises constant strife as each seeks to assert superiority over others, even when this superiority brings no further advantage in terms of natural needs. Humans can bring themselves neither to leave the company of others nor to willingly subordinate their desires to those of their fellows.

For Kant, this “unsocial sociability” is the primary driving force of human progress: “it is this resistance that awakens all the powers of the human being [and] brings him to overcome his propensity to indolence” (8:21). Like Rousseau, Kant suggests that humans’ merely natural needs for food, rest, and sex are sufficiently limited that they do not require excessive development of human capacities. But the capacity to develop new desires, especially in the context of a need to prove oneself superior to others, requires that one cultivate the full range of human capabilities. “Thus happens the first true steps from crudity toward culture . . .; thus all talents come bit by bit to be developed, taste is formed, and even, through progress in enlightenment, a beginning is made toward [forming society] into a *moral whole*” (8:21). At first, this might happen on a purely individual level, as human beings cultivate speed, strength, and dexterity, and then increasingly the ability to imagine and reason, along with the effort to make progress not only in sciences but in the arts. All of these steps are motivated primarily by “ambition, tyranny, and greed” (8:21), but these motives are sufficient to bring people out of indolence and into the hard work of becoming more and more perfect (though not *morally* perfect) human beings. Through humans’ unsocial sociability, nature achieves the great goal of bringing all human natural predispositions to their end. What are at first mere latent potentials for reasoning, character, scientific development, and artistic creativity become actual through nature’s use of humans’ unsocial sociability.

The story does not end with individual progress, however, nor with mere cultivation of individual talents. Kant insists that the ultimate end of nature for human beings includes not merely human *culture*, within which human predispositions are developed, but also a form of society as a *moral whole*: “the greatest problem for the human species, to which

---

<sup>25</sup> Throughout his writings on history, Kant associates Nature with “Providence” on the grounds that “supreme wisdom is required for the fulfillment of this end” (8:310).

*nature compels him, is the achievement of a civil society administering right*” (8:22).<sup>26</sup>

Nature not only seeks to bring about individual cultivation of predispositions and the culture that follows from and facilitates that development; it also aims to bring about just relations among humans, a “society in which freedom under external laws can be encountered . . . , a perfectly just civil constitution” (8:22). This “civil constitution” has two fundamental aspects. First, it involves unions of people under “republican” forms of government.<sup>27</sup> Second, it involves governments at *peace* with each other. True right cannot be established until all nations *together* affirm the principles of right, and this depends upon perpetual peace amongst nations. Human history, then, tends towards a condition within which all human societies will be organized under just, republican forms of government united into a “pacific league” of nations, a “federative union” that can “secure a condition of freedom of states conformably with the idea of the right of nations” (8:356).

Within his moral philosophy, Kant argues that just government and peace among nations are morally required ends for human beings. Given the need for our freedom to express itself in outward deeds, we must establish conditions within which these outward expressions do not do wrong to others. This moral requirement gives Kant a moral reason to believe that history can progress toward such a state:

I shall . . . be allowed to assume that, since the human race is constantly advancing with respect to culture . . . , it is also to be conceived as progressing toward what is better with respect to the moral end of its existence . . . I do not need to prove this presupposition; it is up to its adversary to prove his case. For I rest my case on my innate duty . . . so to influence posterity that it becomes always better. (8:309)

Given the obligation to promote justice in human relations, one is entitled to believe that progress towards this condition is possible. But Kant’s philosophy of history also emphasizes empirical evidence that the moral interest in political right is a real force in human affairs. For example, in the response of the spectators to the French Revolution:

The mode of the thinking of the[se] spectators . . . manifests such a universal yet disinterested sympathy for the players on one side against those on the other, even at the risk that this partiality could become very disadvantageous for them if discovered. Owing to its universality, this mode of thinking demonstrates a character of the human race at large and all at once; owing to its disinterestedness, a moral character of humanity, at least in its predisposition, a character which . . . permits people to hope for progress towards the better. (SF 7:85).

Kant does not argue that the motives of the revolutionaries themselves can be known to be good; ultimate maxims are, after all, inscrutable. But the universality of moral sympathy for those revolutionaries, even without any apparent cause for personal gain, provides empirical

---

<sup>26</sup> The relationship between culture and a just political society shifts through Kant’s works. In the *Critique of Judgment*, a just civil society is part of the ultimate end of nature only as “the formal condition under which nature can achieve . . . the greatest development of the natural predispositions” (5:432). In the “Idea” and *Perpetual Peace*, a just civil society seems to be an ultimate goal of nature for human beings in addition to (and not merely instrumentally towards) the development of human culture.

<sup>27</sup> For Kant, a “republican” form of government is one that works on behalf of the people as a whole in accordance with principles of justice. Such governments can have different structures (monarchic, aristocratic, or democratic).

evidence of a universal moral interest (even if it does not provide evidence that this interest in fact governs most human actions). And this provides a basis for hope that human beings are still capable of genuine moral improvement.

But Kant's primary basis for hope in political progress towards more just civil society is not based on personal duty nor even on the hope that humans' moral interest in justice will outweigh their selfishness. Instead, as in the case of the development of human culture, Kant argues that humans' unsocial sociability itself provides grounds for believing that human societies will ultimately progress towards more and more just institutions.

The problem of establishing a [just] state, no matter how hard it may sound, is *soluble* even for a nation of devils (if only they have understanding), and goes like this: Given a multitude of rational beings all of whom need universal laws for their preservation but each of whom is inclined covertly to exempt himself from them, so to order this multitude and establish their constitution that, although in their private dispositions they strive against one another, these yet so check on another that in their public conduct the result is the same as if they had no such evil dispositions . . . [W]hat the task requires one to know is how [the mechanism of nature] can be put to use in human beings in order to arrange the conflict of their unpeaceable dispositions within a people that they themselves have to constrain one another to submit to coercive law and so bring about a condition of peace in which laws have force. (8:366)

In his *Leviathan*, Thomas Hobbes had shown how the unsociability of human beings – their “greed, diffidence, and pride” – gives rise to human lives that are “solitary, poor, nasty, brutish, and short,”<sup>2</sup> and Hobbes argued that this state of affairs leads humans to subordinate themselves to lawgivers in order to maintain the order necessary for survival. Following this suggestion, Kant argues that even without any moral interests, conflicts among humans will lead them to find laws to which they can subordinate themselves and others in order to achieve the peace and stability necessary for the satisfaction of their desires. Kant further argues, “even if a people were not forced by internal discord to submit to the constraint of public laws, war would still force them from without to do so” (8:365). All that is necessary is *some* people willing to assert their value over others by force in order to bring *all* into communities united under common laws. Moreover, war ends up driving human societies more and more towards republican forms of government (“Idea,” 8:26; PP 8:365-6), and the trials of war (combined with the “spirit of commerce”) drive nations increasingly towards “honorable peace.” As governments become more republican, the people who bear the costs of war increasingly control whether the nation goes to war and thus increasingly seek peace with other nations. And as the benefits of international commerce grow, these benefits outweigh the benefits of war, such that nations seek a peace within which economic exchange can enrich all. Kant concludes by drawing together his empirical argument for political progress with the moral argument in such a way that the burden of the empirical argument is considerably alleviated: “In this way nature guarantees perpetual peace through the mechanism of human inclinations itself, with an assurance that is admittedly not adequate for *predicting* its future (theoretically), but that is still enough for practical purposes and makes it a duty to work toward this (not merely chimerical) end” (8:368).<sup>28</sup>

---

<sup>28</sup> While significant, neither cultural nor political progress exhaust Kant's account of human beings as an historically-progressing species. Kant emphasizes, for example, the importance of pedagogical and

Equally importantly, Kant insists that political progress must be supplemented with the emergence of a “ethical community.”

The dominion of the good principle is not otherwise attainable . . . than through the setting up and the diffusion of a society in accordance with, and for the sake of, laws of virtue . . . Just as the juridical state of nature is a state of war of every human being against every other, so too is the ethical state of nature one in which the good principle, which resides in every human being, is incessantly attacked by the evil which is found in him and in every other as well . . . [This] ethical state of nature is a *public* feuding between the principles of virtue and a state of inner immorality that the natural human being ought to endeavour to leave behind as soon as possible. (R 6:94, 97)

Leaving this ethical state of nature behind requires the establishment of an *ethical* community, just as leaving behind the juridical state of nature requires a political community. But while political community is established by “external legal constraint,” ethical community depends upon mutual encouragement towards virtue; the only “constraint” applicable here is through a supposed divine lawgiver “who knows the . . . most intimate parts of the dispositions of each and everyone and . . . give[s] to each according to the worth of his action” (6:99). Even with the inclusion of God as “moral ruler of the world,” Kant insists that such an ethical community must have “purity: union under no other incentives than moral ones (cleansed of . . . superstition . . .)” (6:102). God’s role as lawgiver is not to provide an additional incentive to good actions, but merely to provide a unified “supreme lawgiver” under whose authority members of that community unite as a single “people” (6:99). As in the case of political and cultural progress, Kant suggests that progress towards this community depends upon the cooperation of nature,<sup>29</sup> but in this case Kant insists particularly strongly that “each must . . . conduct himself as if everything depended upon him. Only on this condition may he hope that a higher wisdom will provide the fulfillment of his well-intentioned effort” (6:101). Whereas political and even cultural progress happens through unsocial sociability, progress towards *ethical* community occurs only in conjunction with properly motivated cooperation.

For Kant, then, human beings are historical beings. Kant insists throughout his works that humans’ historical progress is *not* progress towards happiness: “only culture can be the ultimate end that one has cause to ascribe to nature in regard to the human species (not its own earthly happiness . . .)” (5: 431, see too “Idea,” 8:20). Humans progressively develop innate talents and predispositions, contributing towards a culture within which arts and sciences flourish. We progress towards more just political structures, both within and among states. Educational progress contributes to bringing about enlightenment, a state in which humans think for themselves. And ethical community contributes to moral development. Precisely how far this moral development goes is unclear. Given his transcendental anthropology of desire, according to which each human being is

---

educational progress towards “enlightenment,” “humans’ emergence from their self-incurred immaturity” (8:35), a progress facilitated by political rights – especially freedom of speech (8:37f.) – and also by the emergence and development of better educational institutions (2:445-452; 9:444, 448-51). Kant was, in fact, such an avid supporter of emerging new educational movements that took the striking step of collecting contributions for this project, making himself available for 3 hours a day during the period in which he was most intensely working on the *Critique of Pure Reason* (see 2:452)!

<sup>29</sup> Or “Providence,” see R 6:100-101.

transcendentally free and thereby responsible for her own moral status, Kant seems committed to the view that fundamental moral character is an individual affair. In some of Kant's works on human history, he emphasizes that historical progress is "not . . . an ever increasing quantity of *morality* in its attitudes [but only] . . . an increasing number of actions governed by duty, whatever the particular motive behind these actions may be, . . . i.e. . . . the external phenomena of man's moral nature" (CF, 7: 91). Elsewhere, though, Kant suggests that historical progress does have an effect on human beings at their deepest moral level. The ethical community seems oriented towards making human beings morally good, and Kant suggests that "since the human race is continually progressing in cultural matters (in keeping with its natural purpose), it is also engaged in progressive improvement in relation to the moral end of its existence" (TP, 8:308-309).<sup>30</sup>

One way to think about moral progress in history is in terms of the problem of radical evil discussed in the first part of this chapter. In the context of the *Critique of Judgment's* aim of bridging the gap between nature and freedom, Kant's reference to morally good wills as the final end of nature cannot refer simply to a noumenal will-in-itself, but must also refer to the *expression* of that will in the world. The final end of nature is good human wills actually expressed in concrete human lives. There are several fairly basic ways in which historical progress as the perfection of human capabilities is an important part of good wills actually expressing themselves in the world. Progress in arts and sciences makes it possible for humans who aim for the happiness of others to more effectively promote that happiness, and the good will that seeks its own perfection requires a cultural context within which the resources for that pursuit are available. Moreover, given the necessity of external freedom for the full expression of one's choices, political rights are needed in order for good wills to fully express themselves in the world.

But radical evil poses three deeper problems for the concrete expression of goodness in human lives. Because human beings "started from evil" (6:72), the final end of nature cannot be *perfect* human wills but only wills that unendingly *progress* towards goodness. And given that radical evil involves an ongoing propensity to evil facilitated through self-deception, even this ongoing progress involves struggle against self-wrought evil tendencies. Finally, since human evil is both fundamental and rooted in the human *species*, it is not clear how one could ever begin to progress beyond one's fundamental commitment to prefer happiness to morality.

Kant's account of historical progress can address *at least* the first two issues, and may be able to address the third. We saw in chapter two that Kant postulates immortality as a condition of the possibility of fully satisfying the moral law, but Kant's philosophy of history provides a naturalistic, secular way of understanding immortality. A human life can be considered a good life as a whole insofar as it not only gradually improves in its own individual pursuit of virtue but also works towards an unending progress in the expression of

---

<sup>30</sup> A similar tension arises in Kant's *Critique of Judgment*, where he carefully distinguishes the "final end" of nature in human beings – the good will that is the only unconditional end – from the "ultimately end" of nature in human beings – the political and cultural progress that can be an object of empirical study (see 5:434-6), but then *connects* the ultimate and final purposes of nature, saying that "the ultimate end [is] that which nature is capable of doing in order to prepare him for what he must himself do in order to be a final end" (5:431).

morally good deeds through reforming the society of which it is a part. The historicity of human nature makes it possible for one's own struggle against evil to be part of an enduring struggle of humanity as a species. In particular, and this aligns the first issue with the second, part of one's struggle against radical evil involves enacting social conditions that work to strengthen one's virtue rather than one's evil propensities. The nature of self-deception is such that one seeks both to excuse misdeeds on the grounds of incapacity or comparison with others, and that one seeks to ignore the moral law itself. Progress in culture, for Kant, involves as a central component the "culture of discipline," that "consists in the liberation of the will from the despotism of desires" (5:432). This progress directly combats the frailty of will by virtue of which one lets one's good intentions be overwhelmed by inclinations. And the "ethical community," for Kant, is precisely a community of people constantly reminding one another of their moral obligations, holding one another accountable in a way that, without being judgmental, makes it increasingly difficult to ignore the demands of morality in self-deceptive ways. In the context of human beings as initially radically evil but potentially in revolution against that evil, even not-strictly-moral cultural and political progress can profoundly affect the extent to which one's revolution expresses itself in concrete improvements. Those whose fundamental moral disposition is one of struggle against evil might, in early phases of human history, be largely dominated by their evil tendencies and only show the slightest glimmers of success in the struggle against it, while those at later stages of historical progress, being increasingly armed against the evil principle through social structures that facilitate morality, will express their good wills more and more fully in their concrete, embodied lives.

These sorts of moral progress in history still leave open the question of whether historical progress can go all the way down, actually enabling or facilitating the revolution in one's maxims itself. And here one might take a page from Kant's discussion of supernatural influence. Just as "the concept of a divine concursus is quite appropriate and even necessary" "so that we should never slacken in our striving towards the good" (8:362), but we should not put this appeal to divine cooperation as an excuse for moral complacency; so we might also appeal to moral progress in history as an encouragement that our struggle against evil will bear real fruit in the world, but must appeal to this progress only in such a way that it prevents rather than justifying complacency. Kant's philosophy of history can thereby provide empirical support for the moral hope that is justified religiously by appeal to God's grace and our immortality.<sup>31</sup>

### ***III. Conclusion***

In chapter one, we saw how Kant's transcendental anthropologies of volition and feeling contribute to answering the question "What may I hope?" through the postulates of God and immortality and through the recognition of human beings as ultimate and final end of nature. But when Kant introduced his questions, he associated "What may I hope?" with *religion*, and elsewhere he claimed that *Religion within the Boundaries of Mere Reason* is where he "has tried to complete the third part of his plan," that is, to answer the third question (11:429). While Kant's transcendental anthropology provides an overall framework

---

<sup>31</sup> For further discussion of the possibility of moral progress in history, see Anderson-Gold 2001, Frierson 2003, Kleingeld 1995, Loudon 2000, Wood 1999, and Yovel 1980.

within which hope can be justified, it is only in his religion and history that this framework is given an empirical content. While the empirical anthropology laid out in the last chapter primarily fills in the empirical account of human beings for which Kant's *Critique of Pure Reason* paved the way, the empirical studies of human evil and history are needed to complete his transcendental philosophy as a whole by providing assurance that the empirical world is conformable to the moral demands of freedom.

---

<sup>1</sup> *R* 6:63, see too *R* 6:36-37, "On the Miscarriage of all Philosophical Trials in Theodicy," 8:270.

<sup>2</sup> See *Leviathan*, chapter 13.