

CHAPTER 3: HUMAN EVIL AND HUMAN HISTORY

In the last chapter, we saw Kant's detailed empirical anthropology. While this anthropology does not rise to the level of a "science" in Kant's strict sense, it is a highly systematic account of universal human characteristics. This chapter looks at two further and related aspects of Kant's empirical account of human beings that flesh out Kant's empirical anthropology and complete unfinished business left by the *Critique of Judgment* regarding what we may hope for humanity (see 11:429). First, we look at Kant's account of human evil. For Kant, human beings are "radically" evil "by nature." Despite this apparently glum assessment, however, Kant endorses a realistic hope for human goodness. Second, we 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 progressing towards a future of perpetual peace amongst nations and increasingly cosmopolitan political, ethical, and social lives.

1. 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). Kant's argument for human evil is complicated because of apparently contradictory claims. At times, Kant seems to rule out knowing anything about one's moral status at all, saying that "we can never, even by the strictest examination, completely plumb the depths of the secret incentives of our actions" (4:407; see too 6:36-37, 63; 8:270). But Kant does argue for human evil, and when he does so, he claims both that evil "can only be proved [by] anthropological research" and "experiential demonstrations" (6:25, 35) and also that "the judgment that an agent is an evil human being cannot reliably be based on experience" (6:20). 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 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). From a quick look at these passages, it becomes unclear whether there can even be an argument for human evil, and among Kantians who find such an argument, there is a vibrant debate between those who think that this argument is a priori (e.g. Allison 1990, 2001) and those who think that it is empirical (see Wood 2000:287, 2009 and Frierson 2003).

Fortunately, things are not as hopeless as they seem, and Kant's various statements can be put together into a complicated but plausible anthropological defense of human evil. The key to putting together Kant's argument comes at the beginning of *Religion*:

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 ... 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 "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.

The passage above implies that one can infer maxims from actions. While this might seem to contradict the claim above about the impossibility of self-knowledge, Kant is actually remarkably consistent. Whenever Kant emphasizes the inscrutability of humans' motives, he emphasizes only that we can never know that our maxims are *good*. 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, maxims rather than actions 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). Because these actions are wrong regardless of their ends, one can legitimately infer bad underlying maxims from the performance of such actions. Moreover, because moral inscrutability comes partly from humans' 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 an evil motive, one can reasonably trust that there is no underlying righteous motive. Motivational inscrutability is asymmetrical: one can never know that a person – including oneself – is morally good, but one can know that people are evil.

Even if Kant's claims about inscrutability 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)? Neither experience nor a priori arguments alone are sufficient for Kant's proof of evil. 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. But given evil actions, one knows that if those actions are grounded in freely chosen maxims, then the maxims are evil. So to connect 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's transcendental anthropology has already shown that human actions are phenomenal expressions of noumenal, free choices. In *Religion*, Kant adds an account of the specific structure of the fundamental maxim that grounds evil actions.

In particular, *Religion* 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.

[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 ... 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 only sometimes acts morally never really makes the moral law his ultimate motive, since any law whose application depends upon circumstances cannot be the moral law.

Second, Kant connects his transcendental account of humans' free finitude with his empirical account of human predispositions. 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 to discuss a fundamental "predisposition to the good" that consists of three distinct "elements of the determination of the human being," animality, humanity, and personality (6:26-7). The predisposition to animality includes instincts for self-preservation, sex, and "community with other humans" (6:26). The predisposition to humanity includes innate tendencies to compare ourselves with others and "inclination[s] 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).

By subsuming human volitional predispositions under the general category "predisposition to the good," Kant emphasizes that no natural instincts or inclinations are themselves evil: "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.

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 transcendental perspective on the subordination of moral to nonmoral incentives, no empirical claim can imply anything about human evil: "In freedom alone is evil" (18:212). 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 (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 account of morality from Kant's transcendental anthropology, within which Kant shows both that human beings are transcendently free and morally obligated. 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.

But Kant's argument for human evil is not merely directed towards helping readers recognize evil from-within. He also makes an empirical-anthropological point, that human beings are evil *by nature*. The passage above thus also helps complete Kant's empirical anthropology. 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. And Kant sees empirical evidence suggesting that the empirical character of human volition is structured such that moral grounds are inactive when they conflict with sufficiently strong nonmoral grounds. Kant finds such evidence in the "multitude of woeful examples" of human misdeeds, which shows not only that humans have predispositions that make evil possible but also that we have a volitional structure in which the moral predisposition is made inactive by sufficiently strong sensuous incentives. Given our transcendental freedom (established by Kant's transcendental anthropology), human beings are thus evil. Transcendentally speaking, there is no necessity for human beings to have this volitional structure; it is contingent upon transcendently free choice. But empirically 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, Kant's argument for human evil 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 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 generalize an empirically-grounded anthropological explanatory model. These premises are developed in much greater empirical detail, as we showed in chapter two. The fourth premise is a moral premise, a part of Kant's a priori, transcendental anthropology of volition. The evidence for this claim is thus a priori. If this a priori premise is taken as stipulative, the preliminary conclusion at (5) could be taken as an empirical-anthropological conclusion. That is, if

prioritizing the moral predisposition involves consistency (by definition), 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. In that sense, (5) is an empirical fact. But premise (6) is essentially transcendental; there is no empirical 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 moral evil. The conclusion which is both transcendental 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.

b) The nature of radical evil

Having shown that human beings are evil, Kant elaborates on the nature of evil. Most importantly, Kant emphasizes that human evil is “radical” in that “it corrupts ... the subjective supreme ground of all maxims” (6:37). 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, which can be ordered hierarchically. 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 “whenever I can make use of others to satisfy my needs, I will 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 seek to satisfy some nonmoral desires.” This last maxim is the fundamental guiding maxim of an evil human being, and all other maxims are merely applications to particular cases where 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 clarifies important details about the nature of evil. For one thing, radical evil is not only “itself morally evil, since it must ultimately be sought in a free power of choice” (6:37), but it is also tied to a “natural propensity to evil” that structures particular evil choices that human beings make. Many commentators see this propensity to evil as a *precondition* of radical evil (e.g. 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 evil choices (see Frierson 2003). 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). For one thing, evil cannot be in the human being qua object of empirical investigation but must be traced to the free, noumenal agent that grounds empirically-observable behavior. But even 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 to act from principles, and the way this capacity is used gives empirical evidence of freely-chosen evil.

Kant also describes three ways evil might express itself in one's choices: frailty, impurity, and depravity. The first involves merely a lack of character, an "inability to act according to principle" (25:650). Here the principles of one's higher faculty of desire are good, but when it comes to 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.

The other two forms of evil involve acting in accordance with corrupted principles. "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. Such conditional adherence to the moral law is not real adherence. The final form of radical evil, "depravity," 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). The depraved person might often act in seemingly moral ways, but his 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. 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[, where] our real pleasure consisted in doing something that was forbidden" (Augustine 1961: 47). For Kant, Augustine's self-diagnosis must be mistaken; 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.

Finally, in all of these cases, 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. 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 (see 6:39).

c) 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, because this evil "corrupts the grounds of all maxims" it seems that it

cannot “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). We cannot extirpate evil from our power of choice through that same (evil) power of choice. Radical evil is a consequence of humans’ use of their transcendental freedom. But given that we freely choose evil *as the basis of all of our other choices*, it seems impossible to use that same freedom to rid ourselves of evil. The problem of radical evil is made even worse by the fact that human beings not only choose in evil ways but also cultivate themselves and their environment (especially their social environment) to promote the easy exercise of evil tendencies. Finally, the problem is even more acute because no matter how good one might be able to become, one has chosen badly, so one can never be a person who *always* chooses in accordance with the moral law (6:72). 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 (a *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).

Nonetheless, Kant defends moral hope, the possibility of reforming oneself morally despite one’s radical evil. But this commitment to hope generates a problem: how can one reconcile moral rigorism, radical evil, and moral hope? At one level, Kant does not even try to explain how moral reform is possible given radical evil. He points out that evil cannot be extirpated “through human forces” (R 6:37) and adds, “Some supernatural cooperation is also needed” (R 6:44). This “supernatural cooperation” is ultimately beyond rational comprehension and even practical use (see R 6:117-8, 191; SF 7:43-4). The main role of this “grace” is to reinforce humans’ need to do their part to “make themselves antecedently worthy of receiving it” (R 6:44). Kant emphasizes that the inscrutability of grace is no greater than the inscrutability of freedom and even that humans’ continuing recognition of their moral obligations reveals an enduring “germ of goodness ... that cannot be extirpated or corrupted” (R 6:45-6). 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.

But Kant’s theoretically inadequate discussion of radical evil highlights the proper stance towards the problem.¹ Given his transcendental anthropology of cognition, Kant is 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? If evil is a free choice to subordinate the moral law to nonmoral desires, one must simply subordinate nonmoral desires to the moral law. But radical evil is also a self-wrought tendency to act immorally, and it is, moreover, a tendency evident in humans by nature. And these aspects of radical evil require some 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 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

¹ For discussion of more specifically religious aspects of Kant’s justification of moral hope, see Quinn 1984, 1990; Mariña 1997; Michaelson 1990; Frierson 2003, 2007b, and 2010b.

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 five.

2. Human Beings as a Historical Species

While Kant's conception of human evil draws from and leads to a historical conception of human beings, 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. Nonetheless, during the height of work on his transcendental philosophy, Kant wrote a series of papers on human history that develop his empirical anthropology through, among other things, the claim that human "predispositions ... develop completely only in the species [and over history], but not in the individual" (8:18). The rest of this chapter lays out this historical conception of humanity.

(a) Methodology

Like the anthropology discussed in the last chapter, Kant's historical methodology is primarily empirical. Kant begins his essay "Idea for a Universal History with a Cosmopolitan Purpose" by emphasizing that "human actions" as "appearances ... are determined just as much as every other natural occurrence in accordance with universal laws of nature" and "History. . . concerns itself with the narration of these appearances" (8:17). But history is not "mere empirical groping without a guiding principle" (8:161), and Kant's account of predispositions provides this principle. While the empirical anthropology of the previous chapter focused on predispositions as bases of causal powers, Kant's history studies predispositions teleologically. In his *Critique of Judgment*, Kant argued that organic life could be interpreted via purposive predispositions (5:376). In writings on history, Kant adds that "all natural predispositions of a creature are determined sometime to develop themselves completely and purposively" (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 a purpose for it, which implies that at some point in the normal 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. The full development of human reason in arts, sciences, and politics happens only over 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 develop toward their end. And this assumption provides an "Idea" that can underlie a rationally-guided but empirically-based history of ways humans' natural predispositions unfold over time.

(b) The beginning of human history

Kant's treatment of the earliest human history is laid out in "Conjectures on the Beginning of Human History," which offers a quasi-scientific commentary on the story of humans' creation from Genesis. While some philosophers and anthropologists in the 18th century sought to show how human beings developed from other primates – the issue of the relationship between the upright posture and reason 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, see too 8:179). By "fully formed" Kant means only that humans have all of their natural predispositions, not that these are all fully developed, but even this assumption means that Kant does not 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]" (Herder 2002: 196,189). 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 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 actual freedom.

Kant outlines four steps into actualized human freedom. Blending Genesis with Rousseau, Kant first describes how human beings come to desire objects that are not natural objects of instinct: "Instinct ... allowed ... a few things for nourishment but forbade ... others ... Yet reason soon began to stir and sought, through comparison..., to extend his knowledge of the means of nourishment beyond the limits of instinct" (8:111-12). Humans' cognitive faculties become capable of modifying desires and human beings decide 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.

This first step into freedom is not wholly beneficial. The ability to generate new desires includes an ability to generate unhealthy desires, "desires not only without a natural drive ... but even contrary to it" (8:111). Moreover, freedom over desires causes a new problem, "concerning how he ... should deal with this newly discovered faculty" (8:112). Once 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 framework for making such determinations.

While the first stage in human freedom transformed desires in general, 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. Human beings overlay onto their desire for sexual gratification an interest in the beauty and even personality of the sex object. 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" (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 ... – but also simultaneously it is the most inexhaustible source of cares and worries” (8:113).

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, but they also recognize – albeit obscurely – that every other human being is an “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 think that the earliest human beings had worked out theories of human rights, nor that they actually treated all other human beings as equals. Kant is well aware that human beings seek to dominate each other and treat others as mere instruments. But 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 ... a way of living” (8:112), influence takes a form either of blameworthy domination or of cooperation.

(c) The development of human history

The emergence into freedom marks only the beginning of Kant’s historical anthropology. Before emerging into freedom, human beings were distinguished from animals only by latent predispositions to higher cognitive and volitional faculties. But upon becoming free, humans could become a truly historical species. The claim that “human ... predispositions ... develop completely only in the species” (8:18) comes to the fore, and Kant adds a further claim central to his account of human history: “Nature has willed that the human being should ... participate in no other happiness or perfection than that which he has procured for himself free from instinct through his own reason” (8:19). Humans’ faculties of choosing for themselves generate the structure of human history, according to which all development of human predispositions occurs by humans’ own deliberate work.

Kant almost immediately adds an important caveat to this emphasis on freedom. While human history progresses by means of human choices, Nature uses human choices to achieve ends that diverge from the immediate ends of the choices themselves. In particular, nature uses humans’ “unsocial sociability . . . , i.e., their propensity to enter into society, ... combined with a thoroughgoing resistance that constantly threatens to break up this society” (8:20). 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). Humans’ merely natural needs for food, rest, and sex are sufficiently limited that they do not require much 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. These steps are motivated primarily by “ambition, tyranny, and greed” (8:21), which 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 to fruition what are at first mere latent potentials for reasoning, character, scientific development, and artistic creativity.

The story does not end with individual progress, however. The ultimate end of nature includes not merely 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 nature compels him, is the achievement of a civil society administering right" (8:22). Nature aims for 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" involves unions of people under "republican" forms of government and governments at peace with each other. Human history 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. Thus Kant sometimes "rests [his] case" that that history can progress toward such a state "on [his] innate duty ... so to influence posterity that it becomes always better" (8:309). But Kant's philosophy of history also emphasizes empirical evidence that moral interest in political right is a real force in human affairs. For example, in the response of spectators to the French Revolution, "the mode of thinking of the spectators ... manifests a universal yet disinterested sympathy [that] demonstrates ... a moral character of humanity, at least in its predisposition, which ... permits people to hope for progress towards the better" (SF 7:85). But Kant's primary basis for hope in political progress is not based on humans' moral interests. Instead, as in the case of the development of human culture, Kant argues that humans' unsocial sociability provides grounds for progress towards more and more just institutions. Even "a nation of devils" could solve "the problem of establishing a just state ... in order to arrange the conflict of their unpeaceable dispositions ... so that they themselves ... constrain one another to ... bring about a condition of peace" (8:366). Like Hobbes, 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.

Finally, Kant insists that political progress be supplemented by an ethical community, "a society in accordance with, and for the sake of, laws of virtue" (R 6:94). 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 God as "moral ruler of the world," Kant insists that an ethical community have "purity: union under no other incentives than moral ones (cleansed of ... superstition . . .)" (6:102). As in the case of political and cultural progress, Kant suggests that progress towards this community depends upon the cooperation of nature (6:100-101) but 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.

(d) *Moral progress?*

For Kant, human beings are historical. 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 free and 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* [but only] ... an increasing number of actions governed by duty, ... 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. 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).

One way to think about moral progress in history is in terms of the *Critique of Judgment*'s aim of bridging the gap between nature and freedom. The final end of nature is good human wills actually expressed in concrete human lives. 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 for good wills to fully express themselves in the world.

Radical evil poses 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 one 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 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 strengthen virtue rather than evil propensities. “Ethical community” is a community of people constantly reminding one another of their moral obligations, holding one another accountable in ways 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 evil tendencies and show only 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 will 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 fundamental maxims. And here one might take a clue 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 use appeals to divine cooperation to excuse moral complacency; so we might appeal to moral progress in history as encouragement that our struggle against evil will bear real fruit, but must appeal to this progress only in such a way that it prevents rather than justifies 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.²

SUMMARY

Because of humans’ misdeeds, we can posit a motivational-predispositional structure in human beings that subordinates moral incentives to non-moral ones. Because we are transcendently free, this predispositional structure can be ascribed to moral evil. Thus human beings are evil by nature, and because this evil is “radical,” it seems ineradicable. Still, Kant has hope for human beings. Partly, this hope is tied to the possibility of supernatural grace. But Kant’s hope is also reflected in his historical conception of human beings. Kant’s philosophy of history has three main elements: humans’ emergence from a pre-rational to a rational condition; the development of art, science, culture, and political justice through humans’ unsocial sociability; and the hope for the emergence of an ethical commonwealth for the sake of fostering virtue.

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 claimed that *Religion within the Boundaries of Mere Reason* is where he tried to answer that question (11:429). While Kant’s transcendental anthropology provides an overall framework

² For further discussion of the possibility of moral progress in history, see Anderson-Gold 2001, Cohen 2010, Frierson 2003, Kleingeld 1995, Loudon 2000, Wood 1999, and Yovel 1980.

within which hope can be justified, his religion and history give this framework an empirical content and flesh out his transcendental philosophy by providing assurance that the empirical world is conformable to the moral demands of freedom for radically evil beings like us.

CHAPTER 4: KANT ON HUMAN DIVERSITY

Much of Kant's anthropology emphasizes universality and uniformity. His *transcendental* anthropology implies proper ways of cognizing, acting in, and even feeling about the world that are universally applicable to all people. Even Kant's empirical anthropology describes general properties of human nature; while Kant recognizes that "circumstances of place and time ... produce *habits* which, as is said, are second nature," he insists that anthropology should aim to overcome this "difficulty" in order to "rise to the rank of a formal science" (7:121). And Kant's claim that "the human being is evil by nature" is supposed to be based on "anthropological research that ... justifi[es] us in attributing ... [evil] to human beings" in such a way that "there is no cause for exempting anyone from it" (6:25).

Throughout his life, however, Kant was also preoccupied with human differences. Kant lectured more on "physical geography" than any other subject, and especially during its early years, this course included substantial attention to cataloging differences between different types of human beings. He describes the content of this course in 1765, saying "The comparison of human beings with each other, and the comparison of the human being today with the moral state of the human being in earlier times, furnishes us with a comprehensive map of the human species" (2:312-3). Moreover, from the start of his anthropology course in 1772, Kant included discussion of differences between human beings based on variations in temperament, nationality/ethnicity, and sex. In his published *Anthropology*, Kant emphasizes "an advantage for the reading public" in offering "headings under which this or that observed human quality ... can be subsumed," giving "readers many occasions and invitations to make each particular into a theme of its own, so as to place it in the appropriate category" (7:121-2). Among these "headings" one finds different sorts of talents and inclinations, mental illnesses, temperaments, and ethnic and gender differences.

This chapter focuses on Kant's account of human variation. I start with a brief treatment of individual differences, including mental disorders. I then turn to human temperaments, the four basic affective-volitional structures into which every human being can be classified. Finally, I turn to the two most controversial aspects of Kant's account of diversity, his discussions of sexual and racial/ethnic difference.

1. Individual Variations

Within Kant's empirical anthropology, human beings are unique in their particular configurations of predispositions and powers. Chapter two noted that human beings have universal, natural predispositions that govern cognition, feeling, and desire, but the precise way in which these predispositions unfold is not universal. Many differences between individuals are ascribable to environmental differences, such as why one person plays cricket while another plays baseball or why individuals have different beliefs and tastes. But other differences are, to varying degrees, innate.

The most extreme individual differences are found in Kant's accounts of mental disorders.³ For Kant, mental disorders affect each of the three fundamental human psychological faculties: cognition, feeling, and desire. Because cognition is sub-divided into different powers (imagination, judgment, etc.), Kant distinguishes cognitive disorders according to power is affected and how. For example, dementia (*Wahnsinn*) is "deranged" imagination, while craziness (*Aberwitz*) is deranged reason. Kant also distinguishes mere deficiencies from positive forms of derangement, such that, for instance, stupidity is a deficiency of judgment whereby one simply lacks the ability to figure out whether a particular case falls under a general rule, while insanity (*Wahnwitz*) is a derangement of judgment whereby one groups together disparate particulars under false universals. Further, Kant adds melancholia and hypochondria as cognitive disorders distinct from those that fall under more general groupings. Regarding feeling and desire, Kant treats all disorders of feeling under the general name of "affects" and disorders of desire under the general name of "passions." Both are states wherein a particular feeling or desire overpowers the reflection needed to compare that feeling or desire with others, so a single feeling or desire motivates action without (sufficient) reflection. Finally, Kant describes origins of mental disorders and ways of treating them. Madness is ascribed to a biologically-inherited "germ" that sets on at a particular time and takes on its particular character due to circumstances present when it sets on (7:217). Hypochondria results from a "natural predisposition" (7:104) that has the form of a propensity and can be resisted through "intentional *abstraction*" (7: 212).

While interested in "bringing a systematic division" into mental disorder, Kant also classifies differences between ordinary, mentally-healthy human beings. Such people have the same mental powers that operate by the same general rules, but there are differences in the details of their operation and the relative weight of different influences on thought and action. Kant classifies these human differences into two general categories, those that "indicate what can be made of the human being" and those that "indicate what he is prepared to make of himself" (7:285). The latter is identified by Kant with "character purely and simply" (7:285) and has been discussed in chapter two. The former, including talents, natural aptitudes, and temperaments, involve variations "founded upon ... [different] natural predisposition[s]" (7:286) and describe degrees to which various natural powers are capable of being exercised or improved. Talents refer to "excellence[s] of the cognitive faculty" (7:220), natural aptitude "has more to do ... with feeling" (7:286), and temperament "has ... to do ... with the *faculty of desire*" (7:286). All these natural variations "must ... be distinguished from ... habitual disposition (incurred through habit) because a habitual disposition is not founded upon any natural predisposition but on mere occasional causes" (7:286). Beyond humans' shared mental powers and the differences acquired through different lives and experiences, there are also innate differences in the degrees to which and ways in which mental powers can be exercised.

2. Temperament

Amongst the natural variations that constitute "what nature makes of the human being," the most important is temperament. Whereas talents and natural aptitudes are highly individual, Kant follows a long tradition in holding that one can classify people into precisely four "temperaments": sanguine, choleric, melancholic, and phlegmatic. While these categories were

³ See Frierson 2009a,b and Shell 1996:368-305.

originally developed within ancient medicine, Kant brackets medical and physiological conceptions of temperament (see 7:286). Acknowledging that “temperaments we attribute merely to the soul may have corporeal factors ... as covertly contributing causes” (7:286), Kant divides the four temperaments into sanguine and melancholic “temperaments of feeling” and choleric and phlegmatic “temperaments of activity” (7:286-7, 289). The sanguine “is carefree and of good cheer” and lives in the moment (7:288). The melancholic is serious, thoughtful, and tends towards misanthropy (7:288). The choleric is “*hot-tempered* ... [and] *rash* [and] his ruling passion is ambition” (7:289). Just as the sanguine *feels* quickly and easily but is also quickly distracted, so the choleric *acts* quickly but is quickly appeased. As the sanguine has an excess of (cheerful) feeling, the choleric has an excess of activity. Finally, the phlegmatic has “*lack of emotion*” and “the quality of not being moved easily” (7:289-90).

Although generally dismissed today, Kant’s discussion of temperament is important for several reasons. First, in the absence of some accounts of temperaments, Kant might rightly be accused of failing to recognize the important natural (even biological) differences between human beings. Kant’s doctrine of temperaments, like more contemporary psychological investigations and classifications of human psychological variations, provides his universal and historical anthropology with a necessary supplementary account of human difference. Second, Kant’s account of temperaments is important as part of a specifically *pragmatic* anthropology. I discuss pragmatic anthropology in more detail in the next chapter, but here it is important to note that Kant does not *merely* classify different temperaments. He also emphasizes their characteristics in ways relevant to moral and practical assessments and deliberations. For example, when Kant claims that the sanguine person “makes promises in all honesty, but does not keep his word because he has not reflected deeply enough beforehand” (7:288), his advice not only provides needed warning to the sanguine about their own morally pernicious tendencies but also helps others know how to deal with sanguine companions and even how to properly evaluate the moral status of the sanguine’s broken promises (as flightiness, not deception).⁴ Finally, Kant’s discussion of temperament provides an important “hinge” (Larrimore 2001:270) between Kant’s general and universal anthropology and his discussions of differences between human sexes, nationalities (or ethnic groups), and races. Given the offensive nature of Kant’s views about sex and race, this calls for thinking about what, if anything, distinguishes Kant’s practice of subdividing people according to “temperament” – shared, for example, by those who favor psychological personality tests as a way of improving interpersonal relationships – and his practice of subdividing people according to sex and race.

⁴ In this context, Kant’s note at 15:758-65 on “mistakes” or “failings” (*Fehler*) associated with different temperaments is an excellent example of working through the challenges that those with different temperaments will particularly have to face in living and good and happy life. My reading of Kant here differs importantly from that recently offered by Mark Larrimore. I disagree with Larrimore’s claims that a phlegmatic temperament “is a duty” (Larrimore 2001: 284) and that phlegma should be identified with apathy. Phlegma is a temperament that is *like* apathy and that makes it particularly easy to develop true moral character (which includes apathy), but I take Kant’s repeated claim that phlegma can do what philosophy or wisdom does *without real philosophy or wisdom* to be an important *warning* to the phlegmatic. Whereas a sanguine person who lives life according to principles can and should regard that as a moral accomplishment, the phlegmatic one should not.

3. Differences between the Sexes⁵

Kant's discussion of temperament marks the start of his attempt to classify human beings in terms of generic *types* of human. Throughout his anthropological writings, Kant follows his discussion of human temperaments with an account of "The Character of the Sexes" (7:303). Consistently, and in perfect conformity to feminist characterizations of Western discourse as fundamentally patriarchal, Kant's discussion of differences between men and women focuses exclusively on the unique character of *women*. Kant takes men to be paradigmatic of human beings in general, such that a characterization of "the sexes" involves only showing how the previous characterization of human beings in general must be modified for the "special case" of women.

Kant's *Observations*, his earliest (1764) and most popular anthropological discussion of the sexes, includes both perfect sound-bites of Kantian misogyny – "A woman who has a head full of Greek ... might as well have a beard" (2:229) – and apparent mantras of egalitarianism – "the fair sex has just as much understanding as the male" (2:229). The core of Kant's account of the sexes in *Observations* is that women are primarily beautiful, while men are primarily sublime:

it is not to be understood that woman is lacking noble [sublime] qualities or that the male sex must entirely forego beauties; rather one expects that each sex will unite both, but in such a way that in a woman all other merits should only be united so as to emphasize the character of the *beautiful* ... while by contrast among the male qualities the *sublime* should clearly stand out ... (2:228)

Kant's distinction is both descriptive – women *are* generally more beautiful and men more sublime – and normative: "To this [distinction] must refer all judgments of these two sexes, those of praise as well as those of blame" (2:228) such that "what is most important is that the man become more perfect as a man and the woman as a woman" (2:242-3).

Unless one keeps both the descriptive and normative dimensions of Kant's distinction in mind, Kant's account might seem to preclude virtue in women. Kant says both "It is difficult for me to believe that the fair sex is capable of principles" (2:232, see too 27:49), and "true virtue can only be grafted upon principles" (2:217). This might require, as Jean Rumsey claims, that "women ... are in Kant's view less than ... full moral agents" (Rumsey 1997:131). But such attention to the merely descriptive aspect of Kant's distinction misses Kant's insistence in *Observations* that women *are* capable of virtue, but "The virtue of the woman is a *beautiful virtue*" (2:231). And whereas the principles of which women are incapable "are also extremely rare among the male sex" (2:232), the "love [of] what is good" that serves as the foundation of beautiful virtue is grounded in "goodly and benevolent sentiments" that "providence has implanted ... in [woman's] bosom" (2:232). Whereas few men will attain sublime virtue, women are well equipped for beautiful virtue.

By the time of *Anthropology* (1798), Kant's thought underwent several changes that affect his discussion of women. Some of these reflect Kant's interest in courtship and marriage.

⁵ Throughout his discussion of sex differences, Kant assumes heterosexuality and conflates what we would now call sex-differences and gender-differences.

So, for example, Kant's personal notes and lectures in anthropology increasingly emphasize that it is "an essential condition of nature, that woman must be sought" (25:708), so that "the woman *refuses*, the man *woos*; her surrender is a favor" (7:306). This characterization of "natural" courtship practices contributes to Kant's attention to a womanly "art of appearing" or "art of illusion" (20:61, 69, 121, 140) that provides a way to "govern ... men and use them for their own purposes" (7:304).

Kant also reconceived of the difference between the sexes in terms of an overall natural teleology. From his earliest discussions of women, Kant had referred to their "innate" characteristics (2:229) and insisted that women not only are but ought to be different from men (2:229-30). But his later anthropology takes this further. First, Kant clarifies the extent to which these differences are natural. He recognizes substantial differences between 18th-century European relations between sexes and those in "uncivilized conditions" (7:303), where one sees no significant differences between men and women. But, Kant claims, this greater sameness between men and women is hardly a boon to women, who find themselves, without distinctively feminine sources of strength, in conditions of "domestic animal[s]" (7:304). Consistent with his account of predispositions in general, Kant sees feminine character traits as propensities that require the right conditions to flourish: "culture does not introduce these feminine qualities, it only allows them to develop and become recognizable under favorable conditions" (7:303). Kant also offers specific arguments against those who "dispute this [account of sex differences] in the way one disputes something from the speaker's lectern [to show that they are not inherent to] nature and [they] believe it to be a matter of fashion" (25:709). Kant appeals to "universal and constant" facts about the sexes, including not only that women bear children while men do not, but even that for humans as well as "animals . . . , one sees the female is the refusing, but the male the courting party" (25:709). From such universal characteristics, further characteristics of human females – their abilities to please through illusion, a desire to dominate men through charm, etc – can be explained.

The most important part of Kant's account of women is his treatment of "nature's end in establishing womankind" (7:305). There are two natural purposes for women's distinctive characteristics: the preservation of the species [and] the cultivation and refinement of society (7:305-6). "Nature entrusted to woman's womb its dearest pledge, namely, the species, in the fetus," so women's "fear of physical danger" (7:306) combined an ability to "demand male protection" ensures that the fetus (and thus species) will be threatened neither by excessive boldness on the part of the woman nor physical dangers (against which the man will protect her). The second great end of nature is the cultivation of society: "nature wanted to instill the finer feelings that belong to culture – namely, ... sociability and propriety" (7:306). As we saw in the previous chapter, the development of culture is the great natural end for human beings as a species. There Kant emphasized unsocial sociability as the driving force behind this development. Here he highlights that unsocial sociability has a gendered structure. Men and women are attracted to one another but manifest their superiority in different ways. Women's power over men depends upon increasingly polite and refined social interactions; her direct power is exercised through "modesty and eloquence in speech and expression" (7:306). In order to gain equality, women become adept at social interaction. But as women become more capable of coaxing men, they "claim ... gentle and courteous treatment by the male," who finds himself "fettered ... through his own magnanimity, and led by her, if not to morality itself, to that which is its cloak, moral decency" (7:306, see too 2:241). The apparent

weakness and timidity of women ends up becoming one of the driving forces behind cultural and even proto-moral progress in the human species.

Unfortunately, Kant's increased interest in women as a driving force behind progress in history, even to the point of helping develop a moral decency "which is the preparation for morality" (7:306), was accompanied by profound changes in his overall moral theory, changes that effectively preclude women from virtue. Whereas *Observations* emphasizes the importance of "beautiful" or "adopted" virtues even for men and devotes significant attention to spelling out the details of "the virtue of women" (2:231), Kant's mature moral theory not only does not include, but even seems to exclude, anything that could *genuinely* be called feminine *virtue*. Kant's shift from an empirical and sentimentalist moral theory in the 1760s that allowed different sorts of moral worth based on different aesthetic feelings to a more rigorous rationalist morals in his *Groundwork* and later works that emphasizes a "good will" as the only thing "good without limitation" (4:393) precludes taking seriously as "virtue" anything that does not involve acting out of respect for a pure moral law. Thus when he refers to "feminine virtue" (7:307) in his *Anthropology*, the claim seems to be a mere remnant of an earlier view, a remnant that no longer makes sense in the context of Kant's mature moral theory.

At the same time, Kant's anthropological characterization of women as incapable of male virtue (which becomes the only real virtue) is unchanged. The early claim that "It is difficult for me to believe that the fair sex is capable of principles" (2:232) blossoms into a more technical (and more problematic) claim that certain "feminine principle[s are] hard to unite with a *character* in the narrow sense of the term" (7:308). This "narrow sense" of character is the capacity to act on consistent principles of one's own, a capacity that not only "has an inner worth" of its own (7:293) but is also a necessary condition of a good will (Rumsey 1989, Frierson 2006). Given that woman's distinctive art is an art of appearing, it is perhaps unsurprising that character, which depends upon "not dissembling" (7:294), is unavailable to them. And given that women naturally focus on "pleasing others" (7:305), it is unsurprising that they lack the self-governance required by the moral law.

Given his apparent indifference to treating women as capable of moral worth, Kant's moral and political theory unsurprisingly fail to accord women rights equal men's. Within political theory, women show up the contexts of citizenship and marriage. Kant defines citizens as "the members of a society who are united for giving law" and insists that "all women" are "passive citizens" not "fit to vote" but nonetheless with "freedom and equality *as human beings*" (6:314-5). And whereas in general, society's laws must allow that "anyone can work his way of from this passive condition to an active [citizenship]" (6:315), his insistence that *all* women are passive excludes them from this condition. For Kant, women must remain forever dependent upon husbands for public representation.

In *Anthropology*, Kant adds two important nuances to this account. First, he makes clear that while "woman regardless of age is declared to be immature in civil matters," this is a specifically civil declaration; a wife is, if anything, "over-mature" in her ability to "represent both herself and her husband." But "just as it does not belong to women to go to war, so women cannot personally defend their rights and pursue civil affairs by themselves." Second, Kant suggests that women *do* defend their rights and pursue civil affairs *indirectly*, since "legal immaturity with respect to public transactions makes women all the more powerful in respect to domestic welfare; because here the *right of the weaker* enters in, which the male sex by its nature

already feels called on to respect and defend” (7:209). For Kant, women are excluded from public politics not because of a genuine incapacity, but in order to empower them at home, where they can control their husbands and ensure that husbands take care of the family’s public affairs.

Kant discusses marital rights in the general context of *property* rights, in a section illuminatingly entitled “on rights to persons akin to rights to things” (6:277). There are three ways to acquire “a person akin to a thing,” when “a *man* acquires a *wife*; a *couple* acquires *children*; and a *family* acquires *servants*” (6:227). Of these three, Kant insists that both children and servants are acquired only for a specified period of time, after which they must be granted complete freedom from their parents/masters (6:281, 283). Only women are capable of being “acquired” for life. To be fair, Kant’s account of the husband’s ownership right over the wife is carefully described not as a right to a person *as a thing*, which would blatantly contradict the obligation to respect all others as an end and not a mere means, but only as a right *akin* to rights to things. In fact, Kant’s discussion of the marriage right is one of the clearest places where he articulates the view that women, despite whatever limitations they may have anthropologically, are nonetheless ends in themselves who have a “duty ... to humanity in [their] own person[s]” (6:280). It follows, for Kant, that marriage rights – unlike rights over children and servants – must include entire reciprocity and “equality of possession,” such that just as the husband entirely owns the wife, so the wife in turn entirely owns the husband (6:278). Thereby Kant rules out polygamy, concubinage, and prostitution; and he even insists on “equality in their possession of material goods” (6:278). Consistent with accounts of differences between sexes in *Anthropology* and elsewhere, Kant does not think that husband and wife play identical roles within marriage. In the *Metaphysics of Morals*, where legal rights are at stake, Kant insists that the husband “is to be [the wife’s] master” (6:279). However, consistent with his view that women hold power primarily through charming manipulation of their husbands, while the *law* recognizes the husband as head of household, the wife – through her relational adeptness – “dominates” the husband through his own will.

In the end, Kant’s account of sex difference comes down on the “wrong side” of the most important issues of the day, such as women’s education and citizenship. And while he “gallantly” (e.g. 7:310) praises women’s distinctive charms, his overall account sees them primarily as means to the civil and moral development of *men*. Even Kant’s explicit endorsements of women’s equality (or superiority) fit within an overall attempt to defend and entrench patriarchal political structures. Saying that women have soft power that flourishes in contexts where they lack explicit and formal power is an excellent way to justify denying them political equality. And drawing attention to women’s ability to control men through charm is a good way of discounting the role of rational argument and dialogue at the level of intellectual equals. This discounting can have profound effects not only within marriages – where husbands will expect wives to be charming rather than wise – but also in the education of girls, which education would, for Kant, properly emphasize learning social graces rather than intellectual pursuits (including not only abstract metaphysics but also disciplines that involve more obvious uses of power, like physics, engineering, and politics).

Simply accepting Kant’s views about women is unacceptable not only because they conflict with contemporary assumptions, but also because they conflict with Kant’s own transcendental anthropology. Through all the particular anti-feminist and misogynistic claims in Kant’s eventual account of women’s nature, it is the inability for women in their own right to

have the unconditional worth of a good will that is the most morally and philosophically problematic. Moral responsibility, worthiness to be considered an end-in-itself, and the capacity for a good will are inextricably connected in Kant's transcendental anthropology. Kant explicitly claims that women must be treated as ends-in-themselves (6:278, 280). And there is no evidence that he denies them moral *responsibility*. So Kant needs to provide an account of how women can be capable of moral worth. And doing this will require substantial revisions in his anthropology.

* * *

Today, there are two major and opposing responses to Kant's characterization of women. The dominant response among those sympathetic to Kant "is to say that Kant's views on women are mistaken, that one should instead concentrate on his more important philosophical achievements, and that one can simply leave his theory about the sexes behind" (Kleingeld 1993:140). This response rejects Kant's anthropological characterization of women and extends his general anthropology to include women. Within such an approach, one would affirm – with Kant – that character and rationality are crucial to virtue, but add – against Kant – that women are no less capable of these traits than men. A second response, dominant amongst feminist critics of thinkers like Kant, is to argue that Kant's philosophy as a whole reflects his masculinist bias. Such critics typically agree with Kant's general anthropological claim that there are important differences between men and women, but reject his identification of what is universal and normative for "human beings" with what is universal and normative for men. Carol Gilligan, for example, suggests – with Kant – that for women "morality is conceived in interpersonal terms and goodness is equated with ... pleasing others" rather than in an "understanding of rights and rules" (Gilligan 1982:2). While she rejects the view that this difference is either universal or essential, Gilligan – like the Kant of *Observations* – takes the "different voice" women bring to moral deliberation to be legitimate and needed to balance masculine emphasis on rule-following and personal autonomy (Gilligan 1982, see too Noddings 1984). While Gilligan focuses on differences in moral perspectives, other feminist thinkers have made similar points about Kant's transcendental anthropology more generally. His emphasis on reason and understanding over sensibility has been taken to reflect a masculinist bias in epistemology, one that puts "the Enlightenment conception of a universal, rational subject" above "feminist notions that the self is embedded in social relations, that the self is embodied, and is thus historically specific and partial" (Schott 1997:8).

This divided response to Kant's thought highlights in contemporary form the problem that arises within Kant's own anthropology. Philosophers who find Kant's transcendental anthropology convincing have sought to jettison empirical-anthropological accounts of women than make them seem ill-suited for fulfilling requirements of that transcendental anthropology. But philosophers and empirical psychologists who study sex and gender often end up supporting, if not Kant's specific claims, at least accounts of gender differences that raise similar philosophical problems for a broadly Kantian account of moral and epistemic norms. What seemed to be a tension internal to Kant, one that he lamely resolved by simply settling into misogyny and ignoring the problems this raised for his transcendental anthropology, appears as a real problem for anyone who finds plausible both Kant's arguments for universal norms

governing thought, choice, and feeling and the importance of empirical sensitivity to human differences in thinking about how those norms play out in the world.

One important way to deal with the tension would be through more careful and fine-grained approaches to both transcendental and empirical anthropology. These might find that women and men are not different in ways that have moral relevance. Even if, say, morality requires acting on principles and women tend to be more situational, this does not *necessarily* mean that they lack what is necessary for morality. Morality might *also* require situational sensitivity (see e.g. Herman 1993), and women might also be capable of morally principled action, even if not in the same way or to the same degree as men.

Another middle-ground would accept both Kant's philosophical defense of norms that seem more masculine than feminine and empirical evidence of differences between sexes, but reject the teleological essentialism underlying Kant's explanation of those differences. The recent emphasis on distinguishing "sex" and "gender" obviously fits into this general approach. Insofar as one's gender can be distinguished from one's biological sex, one need not identify biological sex (male/female) with the characteristics associated with particular genders (masculine/feminine), and one can largely ascribe gender characteristics to social factors. As Carol Gilligan emphasizes, "differences arise in a social context where factors of social status and power combine with reproductive biology to shape the experiences of males and females and the relations between the sexes" (Gilligan 1984:2). Kant himself recognizes that the sex differences he discusses arise only in particular social and political contexts (7:304) and his discussion of women's civil inequality (6:314, 7:204) remains at least open to the idea that women's civil immaturity is socially-created rather than natural. Differences between men and women that inhibit women from fully realizing Kantian ideals of autonomy may be due to unjust social conditions that can and should be remedied. This approach could support increased attention to social and political reforms to create a world within which women would have as good a chance as men at measuring up to universally human moral (and other) norms. Kant, unfortunately, rejected this middle-ground in both his anthropology – where he insists that differences between men and women are natural and not merely social (e.g. 25:709) – and in his politics – where his explanation of the passive citizenship of "all women" is combined with both complacency and an account of marriage that seems to reinforce this civil inequality.

Finally, even if it does turn out that there are essential differences between men and women *and* that these differences make it considerably more difficult for women to attain to a good will, one might still – and Kant certainly should – insist that it *is* possible for women to have unconditional moral worth. Modifying Kant's claim from *Religion* (see chapter three), we might say "In spite of [one's sex], 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). In this context, the difference between men and women would be akin to differences between temperaments, not a denial of the *possibility* of virtue for women, but a detailed attention to the fact that women will face greater and different challenges in progress towards virtue than men.

In the end, Kant's treatment of women, in the context of his anthropology as a whole, raises problems and tensions that continue even today to affect thinking about relations between the sexes, and between empirical and transcendental philosophy. But Kant not only rejects the most natural ways of dealing with these problems, but his infatuation with the "charming

difference that nature sought to establish between the two human sexes” (2:228) seems to have made him blindly and complacently unaware of them.

4. Racial and ethnic differences

Sex differences are not the only offensive and provocative part of Kant’s anthropology. When Kant discusses racial and national character he makes statements about other races that are, from our contemporary standpoint, outrageously racist. Moreover, Kant develops a complicated *theory* of race, one that played a role in the development of 19th-century scientific racism and thereby continues to affect the way races are conceived today. Finally, Kant’s moral and political theory at times specifically addresses the relationships between peoples of different racial and ethnic groups. Precisely how these elements fit together is not always clear; and at least the first, and probably the second, share with Kant’s views on women both immediate offensiveness and serious tension with his transcendental philosophy.

Throughout this section, I focus on racial rather than ethnic (or what Kant calls “national”) differences.⁶ Kant was (one of) the first thinker(s) to develop a scientific concept of “race,” and many of his most outrageous comments about other peoples are comments about non-European races. The emphasis of this section on different races is due both to the presence of a systematic race-theory in Kant’s own writings and to the importance of “race” today. For Kant, however, differences amongst Europeans were at least as important as differences between Europeans and others. His published *Anthropology* includes a major section on differences amongst European nations but only two short paragraphs on the character of the races, and even his *Observations on the Feeling of the Beautiful and Sublime* devotes only a “quick look” at “other parts of the world” (2:252) after offering a substantial discussion of differences between European people groups (French, Spanish, English, etc.) (2:243-52).

(a) Kant’s descriptions of other races

Kant’s most disgraceful (published) claims about races are found in his *Observations on the Feeling of the Beautiful and Sublime*. After describing a conversation about the nature of women between an African carpenter and a European missionary, Kant writes of the African’s comments, “There might be something here worth considering, except for the fact that this scoundrel was completely black from head to foot, a distinct proof that what he said was stupid” (2:254-5). Kant’s general characterization of Black Africans in *Observations*, though painfully offensive, is also worth quoting:

The **Negroes** of Africa have by nature no feeling that rises above the ridiculous. Mr. **Hume** challenges anyone to adduce a single example where a Negro has demonstrated talents, and asserts that among the hundreds of thousands of blacks who have been transported

⁶ I also do not discuss Kant’s important comments about Jews (see especially R 6:125ff.); some, but not all, of the present discussion is relevant to assessing these comments. While his views have been seen as important precursors of later German anti-Semitism, they would take the present discussion too far afield because Kant sees the Jews primarily through a cultural and religious rather than racial lens.

elsewhere from their countries, although very many of them have been set free, nevertheless not a single one has ever be found who has accomplished something great in art or science or shown any other praiseworthy quality, while among the whites there are always those who rise up from the lowest rabble and through extraordinary gifts earn respect in the world. So essential is the difference between these two human kinds, and it seems to be just as great with regard to the capacities of mind as it is with respect to color. The religion of fetishes which is widespread among them is perhaps a sort of idolatry, which sinks so deeply into the ridiculous as ever seems to be possible for human nature. A bird's feather, a cow's horn, a shell, or any other common thing, as soon as it is consecrated with some words, is an object of veneration and of invocation in swearing oaths. The blacks are very vain, but in the Negro's way, and so talkative that they must be driven apart from each other by blows. (2:253)

This text does not represent the limit of Kant's offensive comments. In notes from his lecture course on *Anthropology*, Kant claims, in terms reminiscent of his comments about women:

If we compare the character of the Oriental nations with the character of the Europeans, we here thus find an essential difference ... A capacity to act in accordance with concepts and principles is required for character. All Oriental nations are completely incapable of judgment in accordance with concepts ... All Oriental nations are not in the position to explain a single property of morality or of justice through concepts; rather all their morals are based on appearance. (25:655)

And in his *Physical Geography*,⁷ Kant offers a sort of summary of his views of the different races of the world:

Humanity is at its greatest perfection in the race of the whites. The yellow Indians do have a meager talent. The Negroes are far below them and at the lowest point are a part of the American peoples. (*Physical Geography*, 9:316, see too 25:843⁸)

These comments offer but a sample of the dismissive and demeaning views about non-Europeans scattered throughout Kant's writings and lectures.

Kant is not always *entirely* dismissive of other races. In *Observations*, Kant has some admiration for the "sublime cast of mind" of the "**savages ... of North America.**"

⁷ Kant taught a course in Physical Geography for decades (more than any other single course), and, like his lectures on anthropology, student transcripts of these lectures were circulated during Kant's days. Until recently, the main form in which these lectures have been available to Kant scholars has been in an edition put together by Kant's student Friedrich Theodor Rink, likely based on notes from two different courses and included in volume 9 of Kant's works. Volume 26 of Kant's works (published partly in 2009 and partly forthcoming) will include many more versions of student notes from Kant's Physical Geography course.

⁸ The remarks at 25:843, while not wholly reliable because based on lecture notes, are perhaps the most disturbing. There Kant not only sets up an implicit hierarchy of races, but adds that "Negros are not capable of any further civilization ... The Indians and Chinese seem to be static in perfection, for their history books show that they do not know more now than they have long known." (25:843). For both Africans and Asians, Kant alludes to character traits that are not merely physical (civilization and learning) and suggests not only that these races are inferior to whites, but that they cannot ever improve.

Lycurgus probably gave laws to such savages, and if a law-giver were to arise among the six [Native American] nations, one would see a Spartan republic arise in the new world; just as the undertaking of the Argonauts is little different from the military expeditions of these Indians, and **Jason** has nothing over **Attakakullakulla** except the honor of a Greek name. (2:253-4, though cf. 25:1187)

Elsewhere Kant says that “Hindus ... have a strong degree of composure, ... they all look like philosophers, ... [and] they acquire culture in the highest degree” (25:1187). On the whole, however, Kant’s informal “observations” about other races reflect the prejudices of an European satisfied with the superiority of his own race and ready to believe the worst and most degrading claims about others.

* * *

To those who know Kant through his moral philosophy or the universal claims of his transcendental and empirical anthropology, these deeply disdainful comments about other races are disturbing to say the least. How should we respond to comments that seem so out of line with the respect for humanity that Kant emphasizes elsewhere in his work?

Unlike Kant’s comments about sex, where many continue to argue that there are important differences between men and women that may be relevant to moral or epistemic issues, the notion that there are serious innate differences between races that would inhibit members of a particular race from being able to satisfy the demands of Kant’s transcendental anthropology cannot be taken seriously. To those who know Toni Morrison, Jacob Lawrence, Benjamin Banneker, or Wangari Maathai, Kant’s reference to Hume’s claim that “not a single [African] has ever be found who has accomplished something great in art or science” (2:253) would display little more than laughable ignorance if it were not so appalling. With respect race, the issue is not whether to accept Kant’s racial distinctions and adjust his transcendental philosophy or vice versa. No serious thinker today can affirm Kant’s racial observations. But two issues remain: first, whether Kant’s views on non-white races taint the rest of his philosophy such that his claims about, say, moral norms or cultural progress must be abandoned because they are inextricably linked with racism, and second, how Kant – that champion of universal human dignity – could espouse views that seem to deny that dignity to most of the world.

One response to the first issue involves simply dismissing or ignoring Kant’s racially offensive comments. While there are glimmers of Kant’s views of other races in his more well-known writings – e.g. his reference in *Groundwork* to “the South Sea Islanders ... [who] let talents rust and are concerned with devoting life merely to idleness, amusement, [and] procreation” (4:423) – by and large Kant’s best read texts give little explicit indication of his racial views. It seems easy to excise these offensive texts from Kant’s corpus, ignore them, and focus on the parts of Kant’s thought that can and should be candidates for serious consideration today.

The advantages of this approach should be clear. Kant’s philosophy has had profound impacts on metaphysics, epistemology, ethics, aesthetics, philosophy of religion, political theory and so on. Kant’s moral philosophy, in particular, continues to play important roles in safeguarding human rights and individual autonomy. Throwing out Kant’s insights in these areas because of his personal views about other races is a waste. Moreover, Kant himself provides an

important justification for leaving his racially offensive comments behind and focusing on the rest of his anthropology. For Kant, transcendental anthropology (epistemology, ethics, and even aesthetics) must be developed a priori; empirical insights – including observations about human differences – are relevant only later, in thinking about how to apply a priori norms to empirically-situated human beings. Given Kant's own discipline in isolating his transcendental philosophy from his empirical observations, one seems justified in ignoring the latter and reaping the insights of the former.

That said, simply ignoring Kant's comments on race brings important dangers. One danger is that one risks misunderstanding those parts of Kant that one chooses to accept. Kant's apparently off-handed reference to the South Sea Islanders, for example, takes an important stand in 18th-century debates about the moral status of so-called "primitives." Most travelers' accounts, especially of Tahiti (rediscovered by Europeans in 1767), were both "morally provocative" – contrasting Europeans morals with those of the native – and "complimentary," in that travelers to these places typically presented the lives of at least the Tahitians (and often other "savages" as well) as idyllic not only in terms of pleasures but also of morals (Wilson 1998: 317). Kant's insistence on the obligation to cultivate one's perfections (including non-moral talents) is an important part of his moral philosophy, one for which Kant offers an apparently universal justification. But in the context of Kant's claims about other races, this claim can be situated into a general Kantian defense of the superiority of the progressive historical self-conception of an European comparing himself to the rest of the world. In itself, this added insight does not give a reason to *reject* Kant's claim that humans ought to cultivate their talents. But it does force one to look more carefully at the justification and implications of that claim. Precisely how to understand claims that might otherwise seem to be universal, especially when these arise in Kant's transcendental anthropology, cannot be easily settled simply by pointing to Kant's views about races. But paying attention to these views can encourage more nuanced attention to Kant's philosophical views.

This point highlights other risks of simply ignoring Kant's views on race. Kant's claims about other races at least *seem* to conflict with other aspects of Kant's anthropology. At the very least, they raise questions about how Kant could have reconciled his universal anthropology with a view that "Negroes" are irredeemably stupid and "Orientals" incapable of concepts. Trying to figure out how Kant could have held together what seem to be such disparate views can open up new insights into the meaning, limits, and dangers of what might otherwise seem benign aspects of his philosophy (see Eze 1994, Larrimore 1999, and Loudon 2000). Alternatively, showing precisely the way in which, say, Kant's moral theory conflicts with these claims can reveal that moral theory as an important resource for overcoming racism today (Boxill and Hill 2001, Loudon 2000). Finally, failing to pay attention to Kant's negative views on race can mark a missed opportunity to more fully understand the limits of philosophy itself. Investigating how Kant, who claims, "I would feel by far less useful than the common laborer if I did not believe that [my philosophy] could impart a value to all others in order to establish the rights of humanity" (20:44), could hold such offensive views about other races can help reveal some of the causes that continue to prolong racism today.

Despite the dangers of simply *ignoring* Kant's racially offensive comments, however, dismissing Kant's whole philosophy as tainted by racism is worse. For one thing, Kant's transcendental anthropology has played an extremely important role in helping philosophers –

and societies – see racism as unacceptable. Kant’s moral emphasis on respect for each and every person is still one of the most powerful philosophical tools for combating racism. And as we will see in chapter eight, even Kant’s epistemological claim that our world is constructed in terms of imposed a priori categories has helped cultivate awareness of how *nonuniversal* categories of thought can shape experiences, an awareness crucial for cross-cultural understanding. And although Kant’s interest in transforming the study of other cultures into a serious academic discipline was tainted, his conviction that being an educated world citizen requires understanding not only universal characteristics of human beings but also human differences remains an important insight today. Finally, a considerable amount of Kant’s philosophy, at least as that philosophy has been taken up and influences philosophers, can be freed of Kant’s racist views. The mere fact that most readers of Kant – including those whose interpretations are most influential – are virtually (and often completely) unaware of his views on races shows that those views can, at least to a considerable degree, be understood and appreciated independently. As philosophers and scholars continue to explore the implications and impact of Kant’s racial views on his philosophy as a whole, some interpretations of Kant will have to change, and some aspects of his views that might have seemed plausible will now raise more suspicion. But there is at present no reason to think that his philosophy as a whole will need to be dismissed simply because most of his claims about other races must be.

When we turn to how Kant could have held these views, one obvious response – simply ascribing Kant’s racial stereotypes to his eighteenth-century background – is insufficient. On the one hand, his racist views are largely informed by popular prejudices. The contexts in which Kant’s comments occur are almost entirely contexts in which Kant is deliberately seeking to write (or teach) in ways that will be “popular,” “entertaining” for his readers and students (10:146). *Observations*, in which his most outrageous claims occur, was Kant’s most popular book during his lifetime and was written in part to attract students to his lectures. Because negative attitudes towards non-white racial groups were widely shared, Kant’s demeaning comments would likely have enhanced his works’ popularity. His working-class background may have encouraged him to draw divisions between people that would put himself and the upper class students that he needed to attract to his lectures on the same side. Being a relatively poor intellectual who never went more than 90 miles from home, Kant was limited in his data about other races to accounts written by merchants, explorers, and missionaries. In fact, Kant’s courses in *Physical Geography* and *Anthropology* were, at least in part, designed “to make a more certain knowledge of believable travel accounts and to make this into a legitimate course of study” (Wilson 2006:3). Thus Kant was, to a considerable degree, limited by the biases and prejudices of the travel accounts to which he had access and the culture of which he was a part.

On the other hand, however, it is also clear that eighteenth-century thinking about non-white people was not *uniformly* negative. Kant’s empirical sources were often much more generous in their observations than Kant, being written by travelers influenced by a moral ideal of simplicity that seemed well-exhibited in the exotic peoples they observed (Wilson 1998). Among the most important alternative theoretical ways of thinking about races were those of Kant’s own student Herder, who, drawing from similar travel logs and empirical sources, developed a much less patronizing view of non-European nations, and Georg Forster, who not only published travel accounts of his own that emphasized much more positive views of non-Europeans but also specifically criticized Kant’s own race theory as being insufficiently

egalitarian. In both cases, Kant fought *against* more generous portrayals of other races. In his review of Herder's *Ideas*, Kant even wrote:

[F]rom a multiplicity of descriptions of countries one can prove, if one wants to, that Americans, Tibetans, and other genuine Mongolian peoples have no beard, but also, if it suits you better, that all of them are by nature bearded . . . ; that Americans and Negroes are each a race, sunk beneath the remaining of the human species in their mental predispositions, but on the other side by just as apparent records that ... they are to be estimated equal to every other inhabitant of the world; so it remains to the choice of the philosopher whether he wants to assume differences in nature or wants to just everything in accordance with the principle "Everything is as it is with us." (8:62)

This self-awareness about the process of picking and choosing amidst empirical data shows how Kant *could* – even with his sources and his cultural baggage – have developed a different view of non-white races. And Kant should have seen the inconsistency of his dismissal of other races with his own personal and philosophical trajectory. In the end, while one can point to reasons for Kant's demeaning views of other races in his cultural context, Kant's own words require him to acknowledge "the choice of the philosopher." And while one might understand this choice in terms of pleasing the crowd (especially given the public purpose of his *Observations*) or even in terms simply of making his best attempt at getting things right (given the pedagogical purpose of his physical geography and anthropology lectures), one must also admit that Kant, despite his acuity in some areas of philosophy, was neither sufficiently thoughtful nor sufficiently courageous in thinking about other races.

(b) Kant's Race Theory

Kant's most extreme claims about race come in early or informal works, largely disconnected from any formal theory about races. But over a decade after the publication of *Observations*, Kant wrote an essay entitled "Of the difference races of human beings" that marked the beginning of a series of papers in which Kant "invented the concept of race" (Bernasconi 2001:11) or at least became "a leading proponent of the concept of race at a time when its scientific status was still far from secure" (Bernasconi 2002: 146). Kant's racially offensive but informal remarks bear a greater superficial similarity to present-day racism. But his race *theory* arguably played a more significant role in creating conditions for present-day racism by giving "the concept [of race] sufficient definition for subsequent users to believe that they were addressing something whose scientific status could at least be debated" (Bernasconi 2001: 11).

In one respect, Kant's theory of race was deeply *anti-racist*, in that Kant was a staunch defender of "monogenesis" – the view that human beings are a single species derived from a common ancestor – during a period in which polygenesis – the view that, for instance, black Africans and white Europeans are different *species* – was gaining prominence. The different behavior and physical appearance of distant peoples challenged the limits of the European imagination to the point that postulating different species seemed a natural response to the discomfiting possibility that "we" and "they" were the same sort of being. At the same time,

biological science lacked a universally accepted criterion for determining commonality of species.

In response to the growing interest in polygenesis, Kant sought a scientific account of the human species that would reconcile monogenesis with European's desire to distinguish between people with recognizable and heritable differences. The essence of Kant's account is to distinguish "species" from "race." For "species," Kant adopts from his contemporary Buffon what has come to be the standard account in contemporary biology: "animals that produce fertile young with one another ... belong to one and the same physical species" (2:429). Thus, "all human beings on the wide earth belong to one and the same natural species because they consistently beget fertile children with one another" (2:430). Because interfertility does not *necessarily* imply common ancestry, Kant makes his monogenesis explicit: "[H]uman beings belong not merely to one and the same *species*, but also to one *family*, [since otherwise] many local creations [of members of the same species] would have to be assumed – an opinion which needlessly multiplies the number of causes" (2:430).

Having settled on monogenesis, Kant explains heritable diversity with his concept of a race:

Among ... the hereditary differences of animals which belong to a single [species], those which persistently preserve themselves in all transplantings (transpositions to other regions) over prolonged generations among themselves and which also always beget half-breed young in the mixing with other variations of the same [species] are called *racess*. (2:430)

To distinguish racial characteristics from merely environmental differences, Kant insists that racial differences must persist over many generations even after the relevant "race" is transplanted to a different place. Racial differences must also *blend* when members of different races interbreed to distinguish racial differences from what Kant calls "strains" (2:430).

In this way, *Negroes* and *whites*, while not different kinds of human beings . . . , are still two *different races* because each of the two perpetuates itself in all regions and both necessarily beget ... *blends* (mulattoes) with one another. By contrast, *blondes* and *brunettes* are not difference *racess* of whites, because a blond man can have entirely blond children with a brunette woman. (2:431)

As this example suggests, Kant argues that *skin color* "is especially suited" for dividing races (8:94-5). He thus divides humans into "four classificatory differences ... with respect to skin color . . . [:] the *whites*, the *yellow* Indians, the *Negroes*, and the *copper-red* Americans" (8:93).

For Kant, merely defining the concept of race is insufficient. Kant also aims to show how the concept of race can be illuminated by his overall philosophy of biology, and in particular by the role of natural teleology in empirical anthropology. Thus Kant uses his account of natural predispositions to explain both *how* and *why* human beings became differentiated into different races. Starting with the *purpose* of racial differentiation, Kant claims, "The human species was destined for all climates and for every soil; consequently, various germs and natural predispositions had to lie ready in him to be on occasion either unfolded or restrained" (2:435). Different racial characteristics are well-suited for different climates. Given Nature's end – for humans to settle the entire globe –human beings had a variety of predispositions that could

develop in accordance with different local conditions “so that he would become suited to his place” (2:435). For Kant, however, variability with local conditions is not *merely* an ability to adapt; it has a hereditary component:

Once a race ... had established itself ... this race could not be transformed into another one through any influences of the climate. For only the [original representatives of the species] can degenerate⁹ into a race; however, once a race has taken root and suffocated the other germs, it resists all transformation just because the character of the race has then become prevailing in the generative power. (2:442)

What begins as a mere lack of expression of certain natural “germs” (proto-predispositions) becomes, over time, a “suffocation” of those germs. It is unclear whether Kant intends to say that the germs literally die out, or – more likely – that a propensity for them not to express themselves becomes hereditary. In either case, individual adaptations to climate become fixed characteristics. And this, too, has a natural purpose, so that human beings, “over the course of generations ... appear to be ... made for that place” in which they reside (2:435). Nature intends not only for human beings to spread all over the globe, but also for humans to fit well wherever they find themselves. Kant thus charts a middle path between those who claim that human beings are biologically distinct and those that claim that differences are environmental. Differences between human beings are caused by environmental factors, but at least some of these differences become hereditary.

In many respects, Kant’s formal theory of race is less problematic than his informal negative comments about various races. With rare exceptions, Kant’s race essays refrain from describing moral or intellectual qualities as hereditary, and the claim that skin color is necessarily hereditary is not, in itself, particularly offensive. Arguably, Kant’s account of race is even an important step towards a broadly Darwinian account of the possibility of environmentally-caused heritable changes in given populations. But Kant’s race theory raises three new problems for assessing Kant’s philosophy as a whole. First, the stagnation of various natural predispositions within Kant’s race theory raises the stakes of his disparaging comments. If other races have literally lost the *capacity* for moral or intellectual advancement, this poses problems for Kant’s moral theory and philosophy of history that are similar to those raised in the context of the sexes (where women seemed incapable of moral worth). Second, whatever its relationship to his moral theory, Kant’s race theory seems deeply intertwined with the account of natural teleology in his *Critique of Judgment*, which provides the capstone of transcendental anthropology. Even if Kant’s transcendental anthropology could be isolated from Kant’s informal comments about races, it seems harder to isolate his race *theory*. Finally, even if Kant’s race theory is not as immediately offensive as the comments discussed in the last section, by contributing to the development of a scientific, skin-color-based conception of race, Kant arguably played a real historical role in the development of modern racism.

With respect to the first issue, Kant’s works provide mixed evidence about the extent to which he conceived of moral and intellectual attributes as irremediably fixed in races. In

⁹ In 18th-century biology, the technical term “degenerate” lacked many of the negative connotations we associate with it today.

Observations, the early essay with Kant's most atrocious claims about races, Kant adds a crucial footnote:

My intention is not at all to portray the characters of the peoples in detail; rather I will only outline some features ... [O]nly a tolerable level of accuracy can be demanded in such a depiction, ... and no nation is lacking in casts of mind which unite the foremost predominant qualities For this reason the criticism that might occasionally be cast on a people can offend no one, as it is like a ball that one can always hit to his neighbor. (2:243n)

Later, Kant reiterates,

In each people the finest portion contains praiseworthy characters of all sorts, and whoever is affected by one or another criticism will, if he is fine enough, understand it to his advantage, which lies in leaving everyone else to his fate but making an exception of himself. (2:245n)

Here differences between peoples are like those between temperaments, natural advantages or disadvantages that can be overcome. Readers should take negative characterizations of their nation or race as exhortations to moral strength rather than signs of inextricable inferiority. Moreover, although Kant refers to differences between Europeans and Africans as "essential" (2:253), his early comments on racial differences stick to the level of "observations." Kant even claims,

I will not investigate here whether these national differences are contingent and depend upon the times and the type of government, or whether they are connected with a certain necessity with the climate. (2:243n)

Because this section focuses on "The Character of the *Nations*," including non-European races as an afterthought, the footnotes *might* be restricted to Europeans. But they show a way that Kant could reconcile claims about races with his universalist moral theory. And even when Kant develops his formal race theory, in which he argues for essential and hereditary racial characteristics, he insists that "no characteristic property other [than skin color] is *necessarily hereditary*" (8: 94). Thus in anthropology lectures delivered the year Kant published his "Of the different races of human beings," he insists that the "savage Indian or Greenlander ... has the same germs ... as a civilized human being" (25:694). In these and other passages, Kant seems to endorse the view that whatever moral and intellectual differences there are between races can be overcome.

On the other hand, Kant's racial comments often imply heritable and unchangeable moral and intellectual inferiority. Kant's rankings of various races (e.g., 2:441), and his claim that differences between Africans and Europeans are "essential" and "just as great with regard to the capacities of mind as it is with respect to color" (2:253) imply as much. And although his formal race essays emphasize skin color, the last of these essays describes Native Americans as "*incapable* of any culture" (8:176, emphasis added, cf. 10:239), and the first ascribes to them a "half-extinguished life power" while describing Africans as "lazy, soft, and trifling" (2:438). Perhaps the most systematic-sounding claim comes in a footnote of Kant's last race essay. Here, in the context of a discussion of whether African slaves could be used as free laborers, Kant writes:

Should one not conclude ... that in addition to the *faculty* to work, there is also an immediate drive to activity (especially to the sustained activity that one calls industry), which ... is especially interwoven with certain natural predispositions; and that Indians as well as Negroes do not bring any more of this impetus into other climates and pass it on to their offspring than was needed for their preservation in their old motherland ... [where t]he far lesser needs ... demand no greater predispositions to activity. (8:174n)

Strictly speaking, Kant does not claim that Negroes or Indians are *incapable* of activity or industry, but he does suggest a biological basis for the “laziness” ascribed to them in earlier essays, which brings motivational characteristics into the realm of biologically fixed racial differences. In the end, Kant does not strictly commit himself to race differences with moral implications as profound as those of sex, but even insofar as he approaches these sorts of views, Kant’s moral philosophy gives decisive reasons to treat Kant’s characterizations of various races the way Kant suggests in *Observations*, as a catalogue of traits to which individuals can and should make exceptions of themselves. Even this way of reading Kant is hardly without danger. Members of racial groups classified as having particular defects can come either to be demoralized through Kant’s theory or to overcompensate, trying to “prove themselves” in ways that go beyond the actual demands of moral and cultural life. Even interpreted in the most generous way, Kant’s race theory brings problems, and taken too far, it could license the worst racist abuses, as Kant’s seeming support of slavery in the footnote above might forebode.

On the second issue – the relationship between Kant’s race theory and his *Critique of Judgment* – it would be nice to say, as we did with Kant’s informal observations about race, that one can insulate his Critical philosophy from his race theory. In fact, however, as Robert Bernasconi has rightly pointed out, “Kant’s understanding of race is at stake in the discussion of teleology in the *Critique of Judgment*” (Bernasconi 2002:147). But the relationship between race theory, teleology, and the central aspects of Kant’s Critical philosophy (and with it, transcendental anthropology), is not as problematic as some commentators (e.g. Eze 1994) have suggested because it tends to be unidirectional. Race theory supports Kant’s philosophy of biology by showing a particularly interesting way in which it can be applied, but the general points Kant uses in his race theory – such as the distinction between natural history and mere description of nature (2:434n, 8:153f) or the legitimacy of teleological principles in biology (8:157-84) – are, as Kant recognizes, compatible with different empirical accounts depending upon the empirical details to which they are applied. Where Kant’s race theories go astray is not at the level of these general methodological principles but in their specific and misguided application. Thus while problems with the accounts of judgment and teleology in the *Critique of Judgment* could undermine Kant’s race theory, abandoning his race theory altogether, if done for empirical and not methodological reasons, would not jeopardize his *Critique*. Kant’s more general philosophy of biology does not *preclude* scientific racism, which may be an indictment of a sort, but it also does not *imply* it.

The final issue – the role of Kant in the development of modern racism – is complicated. Kant’s defense of monogenesis and emphasis on physical rather than moral or intellectual characteristics is what one might expect a thoughtful, cosmopolitan humanitarian to develop in the 18th century. As a way of making sense of the confusing array of anthropological discoveries faced by Europeans coming into greater contact with the rest of the world, Kant’s race theory might have seemed well-suited to his moral ideals. All human beings are a single species and

hence – one would think – equally worthy of respect. But there are real, hereditary, biological differences between people that manifest in their physical appearance. So far, so good.

But Kant's racial views were not *in fact* limited to physical characteristics. We have seen this in Kant's own life, where he combined a scientific account of race with intensely negative characterization of non-Europeans. But the same held true for those who appropriated the concept of race as a scientific concept in the 19th and 20th centuries. Given the role that "scientific racism" came to play in the entrenchment of racist ideology, especially in Europe and the Americas (see Gould 1981), Kant's role in making this science possible implicates him in those racist ideologies, even if only indirectly. Kant's concept of race was not, of course, the most important influence on racism (even scientific racism) in the 20th century; Darwin was more important. And if Kant's personal views were different, his overall place in the history of racism might well be that of the well-intentioned but naïve humanitarian attempting to combat proto-racist tendencies through science, whose science ended up working to promote racism. But given Kant's personal views about race and the fact that his race theory *did* in fact end up playing a real role in the development of what has come to be modern-day racism, Kant must – unfortunately – be given a prominent place in the history of racism, as he has in the histories of human rights, aesthetics, theology, and other fields where his impact has been more positive.

c) Political issues: Slavery and colonialism

When Kant turns from *descriptions* of different races to *prescriptions* for how to deal with members of other races, the merits of his moral philosophy overwhelm his empirical anthropology. The two most important issues facing Europeans in their interactions with other races were the closely connected issues of colonialism and race-based slavery. And Kant's later political writings involve detailed and impassioned rejections of both. With respect to slavery, Kant's published writings are clear and direct.¹⁰ He refers to West Indian slavery as "the cruelest and most calculated" (8:359) and insists in his *Metaphysics of Morals* not only that no one may sell himself into slavery (6:270) but also that *any* relationship between master and servant can be "at most only for an unspecified time, within which one party may give the other notice" and "children ... are at all times free" (6:283). Whatever forms of indentured servitude might be allowed, the chattel slavery associated with the slave trade and European (especially British) colonialism are excluded.

Regarding colonialism in general, Kant is eloquent:

If one compares [the duty to universal hospitality] with the *inhospitable* behavior of civilized, especially commercial, states in our part of the world, the injustice they show in *visiting* foreign lands and people (which with them is tantamount to conquering them) goes to horrifying lengths. When America, the negro countries, ... and so forth were discovered, they were, to them, countries belonging to no one, since they counted the inhabitants as

¹⁰ Kant's unpublished writings are less clear. In notes written during the 1780s, Kant seems to accept the notion that "Americans and negroes cannot govern themselves. Thus [they] are good only as slaves" (15:878). But Kant's notes for his *Perpetual Peace* include condemnations of slavery – especially the enslavements of non-whites by whites – that are even more impassioned and detailed than those in his published writings (see Kleingeld 2007: 15).

nothing. In the East Indies (Hindustan), they brought in foreign soldiers under the pretext of merely proposing to set up trading posts, but [brought] with them the oppression of the inhabitants, incitement of the various Indian states to widespread wars, famine, rebellions, treachery, and the whole litany of troubles that oppress the human race. (8:358-9)

Kant not only vehemently rejects “counting [non-white peoples] as nothing,” but he shows a surprisingly degree of insight into the deceptive justifications for standard practices of European colonialism. In his *Metaphysics of Morals*, Kant returns to the issue of colonialism. Kant claims that “all nations stand *originally* in a community of land” but emphasizes that this does *not* give anyone the right to possess any land they find. Instead, Kant emphasizes the right of first possession (6:263) and specifically applies this to the case of “newly discovered lands”:

If the settlement [of these lands] is made so far from where [the local] people reside that there is no encroachment on anyone’s use of his land, the right to settle is not open to doubt. But if these people are shepherds or hunters (like the Hottentots, the Tungusi, or most of the American Indian nations) who depend for their sustenance on great open regions, this settlement may not take place by force but only by contract, and indeed by a contract that does not take advantage of the ignorance of those inhabitants with respect to ceding their lands. This is true despite ... it [being] to the world’s advantage ... [A]ll these supposedly good intentions cannot wash away the stain of injustice in the means used for them. (6:353)

Again, Kant not only rejects the practice of seizing land in the “New” World, but specifically addresses two of the main ways in which this practice is justified, through appeals to the greater good and through spurious “contracts.” Despite his personal views about the capacities of non-white peoples, Kant has no tolerance for failing to afford them the rights consistent with our common humanity, and he eloquently defends those rights against encroaching European colonial practices.

There are at least two important ways of reading Kant’s moral and political stances towards other races, both of which partially redeem Kant. First, one might rightly point out that Kant’s moral theory emphasizes the importance of respect for the humanity of others, where humanity primarily involves the mere capacity for choice. Just as Kant insists that a person’s (radical) evil does not justify denying them respect, so he should equally insist that whatever differences there are between Europeans and non-Europeans, as long as non-Europeans are human beings with a capacity for choice, they must be respected. In that sense, Kant’s moral arguments against slavery and colonialism are not only consistent with his overall moral theory but show the power of that moral theory even in the face of extreme personal prejudice. Second, one might find in Kant’s political writings some hint that Kant’s views on other races changed in response to criticisms by Forster, Herder, and others (see Kleingeld 2007). The comments in which Kant is disparaging of other races gradually cease after 1792, such that Kant’s *Anthropology* (published in 1798), while in other respects following the general outline of his *Observations*, refrains from making any mention of his views on race, instead referring readers to a text by another author. This has led at least one scholar to argue that “Kant changed and improved his position[on race] during the 1790s” (Kleingeld 2007:3).

Neither of these ways of reading Kant's later work justifies his racially offensive comments. The fact that in his later years Kant did not use negative characterizations of other races to justify slavery or colonialism does not take away the real harm of depicting them as naturally inferior to whites, and Kant was well aware of the damage to one's personhood that comes from damaging the way that one is perceived in the eyes of others. Even changing his views eventually does not justify Kant's holding them for as long as he did. But just as Kant's other writings on race show the limits of Kant's moral philosophy, these later texts show some of its power. At the very least, the emphasis on universal human dignity in his moral philosophy helped prevent Kant from drawing the worst practical implications of his early views on other races. At best, his philosophy may have even helped him see the errors of those views.

SUMMARY

Kant's anthropology has been criticized for being excessively universal, for painting all humans with the same brush and ignoring the differences between them, for being insufficiently attentive to the idiosyncratic, masculine, or European perspective from which Kant writes. Kant himself was deeply attuned to these concerns. Even before he started developing an a priori transcendental anthropology, Kant insisted on teaching a course in physical geography to "make good [his students'] lack of experience" and equip them to function in the world. An important part of this course was exposing his students to the variety of manners and types of people in the world, and from his early *Observations* to his eventually courses (and book) in anthropology, Kant expanded this part of his geography course to include detailed accounts of human difference.

This attention to human difference as important for being an active world citizen is an admirable aspect of Kant's overall anthropology, but the details of Kant's accounts reveal a darker side to such attention. Kant's resources for understanding human diversity were limited. He never married and his primary knowledge of women was through formal dinner parties and English novels. He never traveled, so his primary knowledge of non-Europeans (and even most Europeans) was from books written by others. Kant made the most of books, voraciously reading travel logs, novels, scientific and medical treatises, and other accounts of human difference. But even with this background, Kant recognized that ultimately "the choice of the philosopher" plays a significant role in how such data is interpreted. And Kant's choices, with respect to both women and non-Europeans, were generally reprehensible. With respect to both women and other races, Kant not only gave isolated disparaging remarks, but also developed theories rooted in his natural teleology that essentialized sexual and racial differences. What could have been a significant *improvement* to Kant's anthropology now serves as a *warning* about the dangers of supplementing a transcendental anthropology with an empirical one, and especially of supplementing a universal anthropology with an account of diversity. Despite its errors and dangers, however, Kant's account of human diversity is an important part of his overall account of human beings, and his insistence that human difference is an important part of what makes us who we are is an insistence that continues to resonate today.