Philosophers' Imprint

volume 24, no. 16
november 2024

Race and Radical Evil:

A New Anthropological Interpretation of Kant's *Religion within the Boundaries of Mere Reason*

Daniel J. Smith

The University of Memphis

DOI: 10.3998/phimp.3059

Kant's 1793 book *Religion within the Boundaries of Mere Reason* has always scandalized readers. Particularly controversial has been the theory of "radical evil" in the book's first two parts, which explains wrongdoing in terms of a "propensity to evil" that is said to be "woven into human nature" (6:30).[1] Kant describes this propensity in terms that appear naturalistic, even biological, as in his insistence that it is "innate" and "universal" among all humans, that it belongs to the "character of the species," and even that it comes from specific "seeds" [*Keime*] that can be located within human anatomy (6:29; 6:30; 6:21; 6:80).

This challenging text has been interpreted in a variety of ways. One prominent approach is an "anthropological interpretation," which purports to take Kant's naturalistic language seriously. First advanced by Sharon Anderson-Gold, this interpretation takes radical evil to be an intrinsically communal phenomenon; it refers to the antisocial elements of human nature that arise in us once we enter society. Allen Wood compares it both to Rousseau's account of *amour propre* and to the concept of "unsocial sociability," taken from Kant's own philosophy of history. In this way, commentators have situated Kant's account of radical evil in the context of his writings on history, politics, and religion, and not just within his moral philosophy.

There is another sense, however, in which Kant's account of radical evil is anthropological, but that has not been part of this debate thus far. This is the surprising fact that the account's leading technical terms were originally developed in the context of Kant's race theory. When Kant writes about a "predisposition to good" [*Anlage zum Guten*] (6:26; see also references to *böse Anlagen* on 6:11 and 6:80n.) and "seeds" [*Keime*] for both goodness and evil (6:20; 6:38; 6:45; 6:26; 6:80), he is using concepts he first introduced to describe the inheritance of racial characteristics. Indeed, this pair of terms, *Keime* and *Anlagen*, forms the conceptual core of his account and stands as his main

1. Kant, Immanuel (1998), *Religion within the Boundaries of Mere Reason*, ed. and trans. Allen Wood and George Di Giovanni, Cambridge: Cambridge University Press. All citations to Kant's work are made to the *Akademie* edition, with translations drawn from the Cambridge edition with occasional modification.

innovation in the field: Sloan refers to Kant's "*Keime-Anlagen* theory" as shorthand.[2] Kant's extensive use of these concepts in *Religion* suggests that the book's argument is indeed anthropological but in a different sense from what has been outlined in the standard anthropological interpretation. His view should be situated not only in relation to his writings on history, politics, and religion but also to his account of the human organism in the essays on race.

Kant's account of radical evil, I will argue, is a theory of organic characteristics that he takes to be rooted in human anatomy. It is like the theory of original sin, not only in that it attributes an innate propensity to evil to every human being after Adam and Eve, but also in that it explains the hereditary transmission of that propensity from generation to generation via sexual reproduction. By citing passages neglected in the existing literature, I will argue that *Religion* also marks a contribution to Kant's natural history of the human species, which includes giving an account of the development of the specific seeds and predispositions that make evil (and good) possible.

Reading Kant's account of radical evil in light of his race theory naturally has implications for how we ought to understand it, which is why I suggest that we need a new anthropological interpretation, rather than just a modification of a previous one. Although I agree with their focus on groups rather than on individual moral agents, I disagree with Anderson-Gold's and Wood's main claim that radical evil is grounded in human sociality. I see it as rooted instead in our organic nature, as something which grows in us as it develops over many generations. In any individual, the strength of the predisposition to good and propensity to evil depends not only on the individual's own activity, I will argue, but also on the reproductive community into which they were born. Although Kant insists that no one can be held morally accountable for the actions of others, he still argues that people's natural predispositions differ based on their varying ancestral lineages. I claim that this idea holds just as much for the predispositions described in *Religion* as it does for those described in his race theory. Individual people will find it more or less difficult to incorporate the moral law into their maxims — and will be more or less tempted to prioritize their own inclinations over it — based on the degree of development of both the predisposition to good and propensity to evil which they have inherited from their forebears.

In Sections One and Two of this paper, I survey previous anthropological interpretations developed by Anderson-Gold, Wood, and Grimm. In Section Three, I introduce my own physiological-anthropological approach, in which the propensity to evil and predisposition to good should be understood as anatomical traits belonging to the human being *qua* organism. In Section Four, I explain Kant's race theory, focusing on the concepts of "seeds" [*Keime*] and "predispositions" [*Anlagen*]; I then show how and why Kant uses these terms to rework the doctrine of original sin. Since he transforms the inheritance of moral guilt into the inheritance of a propensity to evil, passages I cite in Section Five show that Kant sees this propensity as being passed from generation to generation through sexual reproduction. In Section Six, I argue that this natural-historical dimension of Kant's account means that it is implicitly racialized, such that human races are at different levels of moral development based on the degree to which the propensity to evil and predisposition to good have unfolded within their ancestral lines.

1. The Standard Anthropological Interpretation: Anderson-Gold and Wood

Anthropological interpretations of *Religion* stress Kant's own insistence that he is speaking about the radical evil of *human* nature, pointing to passages where he talks about our species as a whole. Anderson-Gold contrasts what she calls the "individual or top-down" perspective of Kant's moral philosophy with a "social" or "bottom-up" perspective.[3]

2. Sloan, Philip (2002), "Preforming the Categories: Eighteenth Century Generation Theory and the Biological Roots of Kant's A Priori," *Journal of the History of Philosophy* 40(2), 251.

3. Anderson-Gold, Sharon (2001), *Unnecessary Evil: History and Moral Progress in the Philosophy of Immanuel Kant*, Albany: SUNY Press, 39.

Concepts such as the "highest good" and the "ethical commonwealth" refer not to atomized moral agents, but rather to the problem of how to integrate the ends of individual actors into a communal or universal whole.[4] *Religion* is an important text for this societal interpretation of Kantian ethics because it deals directly with the relevant components of our nature. It does so most notably in its theory of a "predisposition to humanity" which, Anderson-Gold explains, "represents the basis of the social and cultural dimension of human nature, the dimension through which freedom will historically develop."[5] It is this predisposition to humanity that she thinks is specifically corrupted by the "propensity to evil" [*Hang zum Böse*], which in turns means that the latter is irreducibly social.

Where Anderson-Gold focuses on Kant's positive vision for overcoming evil, Wood's anthropological interpretation emphasizes our present, fallen condition. Wood sees Rousseau as the main influence on *Religion*, arguing that Kant's account is best understood as a rewriting of the *Second Discourse*.[6] Following Anderson-Gold, he also compares Kant's theory of radical evil with "unsocial sociability," and describes these concepts as "different names for the same reality."[7] The key passage is the following, from Kant's "Idea for a Universal History with a Cosmopolitan Intent":

> The means nature employs in order to bring about the development of all its predispositions [*Anlagen*] is their antagonism in society, insofar as the latter is in the end the cause of their lawful order. Here I understand by "antagonism" the *unsocial sociability* of human beings, i.e. their propensity [*Hang*] to enter into society, which, however, is combined with a thoroughgoing resistance that constantly threatens to break up this society. (8:21, emphasis in original)

One can see why commentators find this passage relevant to *Religion*. Its technical terms — "propensity" [*Hang*] and "predisposition" [*Anlagen*] — are also found in the main section headings of part one of *Religion*, as "propensity to evil" and "predisposition to good." These characteristics are said to be components of human nature, but their function in this particular passage is to explain laws of history. This is because Kant "treats history as a branch of *biology*," as Wood explains, and accounts for historical change in terms of "the development of the natural predispositions of the human race as a living species."[8]

Although many have found this comparison illuminating, others have criticized Wood for taking it too far. It is one thing to say that Kant's argument has an anthropological basis; it is quite another to identify that basis fully with the single concept of unsocial sociability, a term that does not even appear in *Religion*. Commentators have charged Wood with confusing levels here, suggesting that unsocial sociability is a mere "symptom," "consequence," or "form of expression" of radical evil, and not a synonym for it.[9]

4. Anderson-Gold, Sharon (1991), "God and Community: An Inquiry into the Religious Implications of the Highest Good," in *Kant's Philosophy of Religion Reconsidered*, eds. P. Rossi and M. Wreen, Bloomington: Indiana University Press, 123.

5. *Unnecessary Evil*, 35.

6. Wood, Allen W. (1999), *Kant's Ethical Thought*, New York: Cambridge University Press, 286; 291–293. Pablo Muchnik takes this idea even further, mapping the details of Kant's account point-for-point onto the *Second Discourse*'s main claims. See Muchnik, Pablo (2009), *An Essay on Kant's Theory of Evil: The Dangers of Self-Love and the Aprioricity of History*, New York: Rowman and Littlefield, 143–149.

7. Anderson-Gold, *Unnecessary Evil*, 4; Wood *Kant's Ethical Thought*, 334; see also Wood, Allen (2020), *Kant and Religion*, New York: Cambridge University Press, 76–79.

8. *Kant's Ethical Thought*, 208.

9. Kemp, Ryan (2015), "The Contingency of Evil: Rethinking the Problem of Universal Evil in Kant's *Religion*," in *Rethinking Kant* vol. 3, ed. Oliver Thorndike, Newcastle: Cambridge Scholars, 110; Grimm, Stephen (2002), "Kant's Argument for Radical Evil," *European Journal of Philosophy* 10(2), 167; Grenberg, Jeanine M. (2010), "Social Dimensions of Kant's Conception of Radical Evil," in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik, Cambridge: Cambridge University Press, 181. This criticism is in Allison,

2. An Alternative Anthropological Approach: Grimm

Stephen Grimm offers an "alternate anthropological approach" that aims to address this issue by situating radical evil at a different level. Drawing on Kant's *Lectures on the Philosophical Doctrine of Religion*, Grimm gives a developmental account of the first appearance of evil in human beings that aims to explain what he calls the "puzzle of universal evil." This puzzle is Kant's surprising and controversial view that all of us have a free choice between good and evil, and yet, without exception, we always choose the latter. Every human being is born with natural inclinations, Grimm notes, and these are present in us before we are mature enough to recognize our moral obligations (elsewhere Kant suggests that this maturation usually happens between the ages of 8 and 10 years).[10] Kant sees nothing wrong with inclinations taken in themselves, and even argues this point against the Stoics at 6:58. Once reason emerges, however, the inclinations lose their original innocence, and thus take on a new role as potential competitors to the moral law as motivators for our action. This status does not make inclinations intrinsically evil, but it does transform them into the *occasion* for evil, which consists in "reversing the moral order" of our incentives, and therefore giving inclinations and natural predispositions priority over the moral law (6:36). This reversal is not possible before reason emerges, which is why Kant compares his view to the Pauline idea that "sin follows upon the law" (28:1079, see Romans 7:7).

This account of human moral development, Grimm suggests, explains why Kant can say that all human beings have chosen evil, and that this choice was made freely and is therefore imputable. By the time reason develops, "our natural needs and inclinations have already acquired a kind of foothold within the soul which reason finds nearly impossible to dislodge."[11] It is always possible to order one's incentives correctly from the very beginning, and to maintain the moral law in first rank. But in practice that will be extraordinarily difficult — so difficult, in fact, that it may never have been done. Kant's strict "rigorism," combined with the fact that this rank ordering is purely rational and therefore outside time, means that one need yield to temptation only once in one's life to reveal that one's most fundamental priority was not the moral law all along. We all experience this same developmental process of becoming habituated to various natural inclinations before our reason develops, and so Kant can say that evil is universal for human beings, though without being necessary. In this sense, our choice for evil is contingent but universal among humans, which is just what Kant needs it to be in order to preserve freedom.[12] For Grimm, then, and contra Anderson-Gold and Wood, the root of human wrongdoing is not our social nature, but our "composite nature" — as both animal beings with natural predispositions, and moral beings able to grasp the moral law through reason. Without the former there would be nothing to counteract the moral law's motive force; without the latter, we would remain in a state of pre-moral innocence. Were we not intrinsically composite beings subject to the demands of both nature and freedom, there could be no such thing as radical evil.

Another advantage of Grimm's approach, one he does not note, is that it enables him to address another problem with Anderson-Gold

Henry (2021), *Kant's Conception of Freedom: A Developmental and Critical Analysis*, Cambridge: Cambridge University Press, 471.

10. See 5:155 and 8:286, cited in Pasternack, Lawrence R. (2014), *Kant on Religion within the Boundaries of Mere Reason*, London: Routledge, 110.

11. Grimm, "Kant's Argument," 171.

12. In an infamous passage, Kant refers to the need for "anthropological research" (6:25) to justify attributing either a good or an evil character to the human being. Many see Kant here as hopelessly confusing the empirical and the transcendental, as it is obviously impossible to confirm by observation that every human is evil. For my part, I read Kant's reference to "anthropological research" as describing a possible *dis*confirmation of his claim about the contingent universality of evil, by investigating whether any morally pure "saints" have existed (see the somewhat similar argument in Palmquist, Stephen (2008), "Kant's Quasi-Transcendental Argument for a Necessary and Universal Evil Propensity in Human Nature," *The Southern Journal of Philosophy*, XLVI, 263–4). Kant suspects that no such person has existed but is open to correction if anthropological research shows otherwise, since saints are at least conceptually possible. Part two of *Religion* concerns the Son of God, who is certainly one candidate, but there Kant is speaking of a pure idea of practical reason and leaves undecided the empirical question of whether he existed as a real historical person.

and Wood's interpretation: their view appears to overlook certain important elements of Kant's argument. Allison has brought out well the many contributions *Religion* makes to Kant's late moral philosophy, including the "incorporation thesis," the distinction between legislative *Wille* and executive *Willkür*, and his "rigoristic" account of our "disposition" [*Gesinnung*], which must have the character of either good or evil.[13] These ideas are all non-anthropological concepts prominent in *Religion*, but they play no role in standard anthropological interpretations. The text concerns anthropology, as its naturalistic language suggests, but also deals with *a priori* moral philosophy; any convincing interpretation must account for both elements. Grimm focuses on the anthropological side, but his reference to our "composite nature" as the source of radical evil opens a place for *a priori* elements of Kant's analysis to sit alongside anthropological ones, without reducing either dimension to the other. Although I shall give a different account than Grimm of the anthropological side of Kant's account, I agree with him that radical evil comes from the jointure of the moral and the anthropological.

3. A Physiological–Anthropological Interpretation

As I read it, then, Kant's theory of radical evil has two parts that correspond to these two sides of our composite nature, which must not be confused since they arise from different standpoints. The first is a conceptual investigation into the conditions for the possibility of accountable wrongdoing, as Allison maintains. Because Kant believes that "evil originates from freedom" (6:45), he cannot conceive natural inclinations as the basis of evil, as he arguably did in the *Groundwork*, even if they are still the means by which it is expressed in us. To bridge the gap between nature and freedom, Kant must explain how we freely adopt or "incorporate" incentives flowing from inclinations into our maxims. This he does in his account of the original adoption of a "disposition" [*Gesinnung*], the meta-maxim that determines which maxims we will adopt in the first place, and thereby which incentives we let influence our action in general. This part of Kant's theory is controversial, to say the least, and has seen many interpretations.[14] Post-Kantian philosophers such as Fichte and Schelling found this part of the theory especially interesting but ultimately unconvincing within Kant's own framework, and revised it within their own accounts of evil.[15] I have explored some of these issues elsewhere, and so will not discuss them here.[16] What matters for now is only that Kant's theory of "disposition" is supposed to explain how inclinations, which originate in our organic nature, can be taken up or "incorporated" into our maxims in a transcendentally free act that can be imputed to us.

The second part of Kant's account is an essentially anthropological investigation into the nature of the subject that does the incorporating: namely, a human being with the range of propensities and predispositions attributed to us in *Religion*. Without a predisposition to the good, we would not be capable of following the moral law in the first place, and so we could not be held responsible for transgressing it. Without a propensity to evil, though, we would not be able to explain the apostle John's lament that "the world lieth in evil," and the many confirmations of this judgment by poets, historians, and other religious traditions (6:19). The propensity to evil is Kant's name for the natural tendency to prioritize inclinations above the moral law in their ongoing struggle for influence over our actions. Although we are all afflicted

13. See Allison, Henry (1990), *Kant's Theory of Freedom*, New York: Cambridge University Press, especially chapters 7 and 8, but see also (2001), "Ethics, Evil, and Anthropology in Kant: Remarks on Allen Wood's *Kant's Ethical Thought*," *Ethics* 111, 594–613 and *Kant's Conception of Freedom*, chapter 10.

14. See, for example, Allison, "Ethics, Evil, and Anthropology," 607–608; Peters, Julia (2018), "Kant's *Gesinnung*," *Journal of the History of Philosophy* 56(3), 497–518; Lawrence, Joseph (2002), "Radical Evil and Kant's Turn to Religion," *The Journal of Value Inquiry* 36, 319–335.

15. On the reworking of radical evil in Fichte's *System of Ethics*, see Ware, Owen (2015), "Agency and Evil in Fichte's Ethics," *Philosopher's Imprint* 15(11), 1–21; on the same in Schelling's *Freiheitsschrift*, see Alderwick, Charlotte (2015), "Atemporal Essence and Existential Freedom in Schelling," *British Journal for the History of Philosophy* 23(1), 115–137.

16. Smith, Daniel J., (2021), "An Ethics of Temptation: Schelling's Contribution to the Freedom Controversy," *European Journal of Philosophy* 29(4), 731–745.

by this propensity to evil, which makes it difficult to order our incentives correctly, we are not compelled to act on its basis, because we also have a countervailing "predisposition to personality" that makes us susceptible to "respect for the moral law *as of itself a sufficient incentive to the power of choice* [*Willkür*]" (6:27). So long as this predisposition also develops within us, it remains possible to follow the moral law as we ought. We are thus not doomed to vice and sin. Pasternack draws a helpful distinction here, between "the facticity of the propensity to evil" and "what we do with that fact."[17] Kant preserves the conceptual possibility for every individual to resist this corrupting influence, even though it is deeply rooted in our nature.

I parse out this distinction in terms of *Religion*'s main concepts. When Kant describes dispositions, free will, maxims, incentives, and so on, he speaks the language of his *a priori* moral philosophy. But when he describes propensities and predispositions, human nature, and seeds for both good and evil, he speaks the language of anthropology. "Anthropology" does not refer merely to empirical observations or generalizations about human behavior, as some have suggested in this context.[18] In the preface to the *Anthropology*, Kant distinguishes between "physiological anthropology" and "pragmatic anthropology," with the former describing "what *nature* makes of the human being" and the latter "what *he* as a free-acting being makes of himself, or can and should make of himself" (7:119).[19] Where previous anthropological interpretations have focused on pragmatic anthropology, my argument is that *Religion* also marks a significant contribution to *physiological* anthropology—that is, it describes the human raw material that Kant's practical philosophy has to work with.

17. *Kant on Religion*, 108. Of the commentators cited here, Pasternack's account of the propensity to evil seems to me closest to my own.

18. See e.g. Quinn, Philip (1988), "In Adam's Fall, We Sinned All," *Philosophical Topics* 16, 111; Guyer, Paul (1998), "Immanuel Kant," in *The Routledge Encyclopedia of Philosophy*, ed. Edward Craig, New York: Routledge, 192.

19. See the helpful discussion of this distinction in Lu-Adler, Huaping (2023), *Kant, Race, and Racism: Views from Somewhere*, New York: Oxford University Press, 55.

Before introducing concepts from Kant's race theory, we must note that this complex jointure of the anthropological and the properly moral has long been a focus of the literature on Kant's account of race and especially his rac*ism*.[20] Although Kant famously holds that pure moral philosophy must be "completely cleansed of anything that might be only empirical and that belongs to anthropology" (4:389), Kant's essays on race nonetheless describe differences between human races in terms of moral characteristics as well as physical ones. As such, many have wondered how to reconcile these apparently conflicting ideas.[21] One influential approach has been to treat the tension as a problem of scope. The issue, according to Mills, is not that Kant's moral claims fail to be universal in their form, but rather that he did not see the entire global population of "biological humans" as subjects of those claims in the same way.[22] Mills asks what the persons Kant describes in his moral philosophy are like, which is different from the question of how those agents ought to act once identified. The first question is anthropological, although it is relevant to moral philosophy because it determines the scope of its application.

20. This issue is tackled in Mensch, Jennifer (2017), "Caught between Character and Race: 'Temperament' in Kant's Lectures in Anthropology," *Australian Feminist Law Journal* 43(1), 125–144. But variations of this idea have been explored in much of the literature beginning with Eze, Emmanuel Chukwudi (1995), "The Color of Reason: The Idea of 'Race' in Kant's Anthropology," *The Bucknell Review* 38(2), 200–241; Bernasconi, Robert (2003), "Will the Real Kant Stand Up," *Radical Philosophy* 117, 13–22; Mills, Charles (2005), "Kant's *Untermenschen*," in *Race and Racism in Modern Philosophy*, ed. Andrew Valls, Ithaca: Cornell University Press, 169–193; Huseyinzadegan, Dilek (2019), *Kant's Nonideal Theory of Politics*, Evanston: Northwestern University Press; and more recently Huseyinzadegan, Dilek (2022), "Charles Mills' 'Black Radical Kantianism' as a Plot Twist for Kant Studies," *Kantian Review* 27(4), 651–665; Shorter-Bourhanou, Jameliah (2022), "Reinventing Kant?", *Kantian Review* 27(4), 529–540; Lu-Adler, *Kant, Race, and Racism*.

21. See, e.g. Mills, "Kant's *Untermenschen*"; Hill, Thomas and Boxhill, Bernard (2001), "Kant and Race," in *Race and Racism*, ed. Bernard Boxhill, Oxford, Oxford University Press, 448–471; Fleischacker, Samuel (2023), "Once More Unto the Breach: Kant and Race," *The Southern Journal of Philosophy* 61(1), 3–28; Lu-Adler, *Kant, Race, and Racism*, chapter 1.

22. Mills, Charles (2014), "Kant and Race, Redux," *Graduate Faculty Philosophy Journal* 35 (1–2), 126.

Here, I shall not address the question of exactly how Kant himself resolves this issue. My argument is just that, in *Religion*, Kant also asks a physiological–anthropological question, and uses the same conceptual tools to answer it. What must human beings be like, given that we are both responsible for our actions and beset by a radical tendency to violate the demands this responsibility places on us? What propensities, predispositions, and seeds must we ascribe to the human organism to make sense of our sorry but ultimately hopeful condition? This question, about the anthropological conditions for the possibility of moral agency is again brought out well by Grimm's description of our "composite nature" as the source of radical evil. We could imagine a different race of rational beings that did not have natural predispositions at all, or one that lacked the capacity to grasp the moral law (I leave open the question of whether Kant thought of any human races in these terms). In neither case would the conditions hold for those beings to be accountable moral agents like us. Although Kant's moral philosophy is formally universal, *a priori*, and cleansed of everything anthropological, it nonetheless requires as its subjects beings with certain characteristics. While the issue of what these beings must be like is controversial for Kant's race theory, I argue that it is also relevant to his account of radical evil, which is a contribution not only to pragmatic anthropology but also to physiological anthropology.

4. Seeds and Predispositions in Kant's Race Theory

In the 1775 essay "Of the Different Human Races," Kant's goal is to define a new stratum of classification for living beings. Although significant differences exist between human beings from different parts of the world, all groups seem able to successfully reproduce with all others, which proves through Buffon's rule that there is a unified human species. Earlier race theorists had invoked geographical, climactic, and environmental factors as the causes explaining the differences between races, but Kant knew that these were at most partial solutions. If climate were the only factor, for example, then people's race would change as they migrated across the globe. One need only think of the United States, but Kant's own example is the Romani people, whom he believes are descended from India but had lived in Europe for three hundred years without any changes to their skin color (8:105).

Kant's approach to this newly defined stratum of race is to distinguish between different types of inherited characteristics. Many non-racial characteristics are also transferred by heredity, but Kant's argument is that the laws governing their inheritance differ from those associated with race proper. How does one distinguish a racial hereditary difference, like skin color, from some non-racial hereditary difference, like hair color? If a blonde and a brunette reproduce, their child will usually have either blonde or brunette hair. By contrast, Kant argues that if a white person and a black person produce a child, the child's skin tone will necessarily lie halfway between that of the parents, making such children what he calls "half-breed children or *mongrels* [*Blendlinge*] (mulattos)" (2:430)—or, in the even more charged and demeaning language of the second essay, "whites who have bastardized [*verbastert*] with blacks" (8:105).

For Kant, any theory of race must account for these two things: first, it must allow that racial differences are preserved as people migrate; second, the theory must follow what the second essay formally names "the law of necessary half-breed generation" (8:95). His solution is to posit the existence of "seeds" [*Keime*] and "predispositions" [*Anlagen*] as the causes of particular organic characteristics. These "purposive causes" are opposed to merely physical causes, like air and sun. Kant gives the example of a bird species that can adapt to life in both hot and cold environments. If one of those birds moves from a hot climate to a cold climate, it has the remarkable ability to grow an extra layer of feathers to keep it warm. This special capacity, Kant thinks, can be explained only by something pre-formed lying dormant within the organism until the requisite occasional causes affect it. The birds are born not only with an ordinary layer of feathers for their hot climate, but also with "seeds" for a new layer that will sprout under specific conditions. Where "seeds" refer to individual traits, "predispositions" refer to "the size or relation of the parts to one another" (2:434),

although in practice Kant generally refers to seeds and predispositions interchangeably.

The final relevant component of Kant's *Keime-Anlagen* theory is what Zammito has termed its "irreversibility thesis."[23] Once the seeds begin to germinate, their development cannot be stopped. The bird that acquires an extra layer of feathers in the cold will retain that extra layer, even when it returns to a hot climate. As a seed sprouts within a particular part of the organism, it "suffocates" (2:442) all other seeds related to the same part, which means the disappearance of all the other potentials that the organism would have had to be well-adapted to living in other conditions. These seeds and predispositions are then transferred to the next generation *as developed*; the growth does not begin anew in each individual (Kant wrote long before the familiar Darwinian critique of the inheritance of acquired characteristics). In other words, if the bird were to reproduce once it had a new layer of feathers, its offspring would inherit those feathers from their moment of birth — even though the original occasioning causes were never present, and even though they may live their whole lives in a hot climate where the extra feathers are not needed. For this reason, "seeds" and "predispositions" unfold over large timescales and at the level of a species or race, and not at the level of the individual. This is why Kant speaks of a gradual development of the predispositions of humanity as a whole taking place over many generations. Outside the race essays, these proto-biological terms appear most prominently in Kant's philosophy of history, which is also concerned with longer time scales (recall Wood's explanation that for Kant, "history is a branch of biology").

This theory of hereditary characteristics therefore also implies a specific natural history for the human species. Nature originally designed human beings to be able to live in every climate. The first human "lineal stem formation" — Adam and Eve — must have been created with all the preformed seeds and predispositions belonging to humankind as a whole. But as generations passed and their offspring spread across the earth, different climates caused different seeds to germinate. The germinating seeds allowed each group to adapt to their local environment, but snuffed out all other possibilities they would have had to adapt to others. Kant divides the earth into four climates and posits four corresponding fundamental races of human being, each well adapted to its ancestral home. The environment once played a role in developing those characteristics, but it does not do so anymore; now, all racial characteristics are fixed and permanent. Mixed race peoples exist, but it is clear that Kant sees their existence as a disturbance of nature's plan, as offspring become badly adapted to both climates (2:431).[24] Kant thinks that race-mixing is not beneficial to the species because, according to his theory, it leads to the extinction of characteristics without any new ones being produced (7:320); he worries that too much "interbreeding" could cause currently existing races, like the white race, to "go extinct" (8:105–6).

Why would Kant introduce concepts from a proto-biological theory of heredity into the apparently far-removed context of a book about rational religion? At first sight these discourses seem to have little to do with each other. The connection becomes more understandable, however, when we remember that Kant's account of radical evil is a re-working of the Christian doctrine of original sin. This doctrine is in part a moral theory, and is thus supposed to explain our fallen nature and perhaps what we ought to do about it. But the theory has always included an anthropological part as well, which explains how this defect in human nature is passed on through the generations via sexual reproduction. It is this element of Kant's theory that previous anthropological interpretations have overlooked, because they do not see *Religion* as contributing to natural history or physiological anthropology.

23. See Zammito, John (2017), *The Gestation of German Biology*, Chicago: Chicago University Press, 238, where the idea is described as "one of Kant's decisive interventions."

24. On the status of mixed-race people in Kant, see Gorkom, Joris van (2020), "Immanuel Kant on Race Mixing: The Gypsies, the Black Portuguese, and the Jews on St. Thomas," *Journal of the History of Ideas* 81(3), 408–9; Bernasconi, Robert (2011), "Kant's Third Thoughts on Race," in *Reading Kant's Geography*, eds. Stuart Elden and Eduardo Mendieta, Albany: SUNY Press, 298–299.

One might object to my interpretation that, in a well-known passage, Kant specifically rejects the idea of a hereditary transmission of evil as the "most inappropriate" way of representing the propagation of moral evil through the human species (6:40). As the line he quotes from Ovid implies, one cannot attribute guilt to an individual for an act that that person did not freely carry out themselves. But this is why I insist so strongly on distinguishing the moral and the anthropological standpoints: I am not arguing that we inherit *moral evil* from our kin. Kant distinguishes between the hereditary transmission of guilt — a view he adamantly opposes as ruinous to morality — and the hereditary transmission of a mere *propensity* to evil. Like an addiction (6:29n.), this propensity actively solicits us to transgression, making it harder to do what we should, but it poses no threat to the imputability of our actions, because we always have the possibility to choose to follow the moral law, which remains a sufficient incentive.

5. A Natural History of the Propensity to Evil: Jesus, Mary, and Adam

I have been arguing that *Religion* makes an as-yet unrecognized contribution to Kant's natural history of the human species, focused on the propensity to evil as a hereditary characteristic.[25] Next, I will analyze two passages that identify important moments in that natural history, which are also based on Church doctrines: Kant's account of the exceptional human being who did not have a propensity to evil because of his virgin birth, and Kant's account of the origin of evil in the very first human being.

In a long footnote (6:80n.), Kant addresses the question of how the Son of God could have been born without an innate propensity to evil. It may seem like an obscure theological problem, but the issue is important to Kant because he needs the Son of God to be similar enough to us physiologically that he can serve as a "prototype" for the ideal of moral perfection. If the Son of God were conceived as fully divine — as an angel, say — then his moral purity would be "innate," and we would not think that we could emulate him; he would fail in his function as a role model. For this reason, he must be understood as a human being like us, "afflicted by just the same needs and hence also the same sufferings, by just the same natural inclinations" (6:64). Yet, although he has inclinations, he does not have a radical tendency to prioritize them over the moral law. The idea of a virgin birth suits Kant's purposes perfectly because it explains the Son of God's generation as having a human component, which explains why he is like us, and a divine component, which explains why he is not radically evil.

While we can be sure that the divine father did not pass on the evil seeds, the human mother remains a problem. Kant did not believe in Mary's immaculate conception, so he needs a theory of her supernatural pregnancy to explain why she did not pass on the seeds for a propensity to evil that she must have inherited from *her* human parents. Such a theory, Kant explains, will be "on one side merely sensible, but on the other nevertheless moral, hence intellectual" (6:80n.). It is merely sensible in that it aims to explain the means by which a particular organic trait passes from one individual to another, but it is also moral and hence intellectual because it arises not from empirical observation, but from a consideration of what is required conceptually for the possibility of a purely practical idea that "resides in our morally-legislative reason" (6:62). We know on *a priori* grounds that our practical reason must include an idea of the Son of God as a prototype, and so when we turn to anthropological matters, such as how he could have been born without a propensity to evil, the theory we end up with must not contradict this fact.

This is a complex conceptual problem — here is Kant's characteristically ingenious solution:

> We would have to assume the theory that the seeds of the descendants *pre-exist* in the progenitors, not, however,

25. It is striking that in the edited volume *Kant's Anatomy of Evil*, no contributor refers to "anatomy" or develops an interpretation taking anatomy to be a relevant part of Kant's account of evil (Anderson-Gold, Sharon and Muchnik, Pablo eds. (2010), *Kant's Anatomy of Evil*, Cambridge: Cambridge University Press). If that title had not already been used, the present paper could have been called "Kant's Anatomy of Evil."

> the theory that these seeds develop on the *female* side (for then the consequence is not escaped) but on the *male* side alone (not on the part of the *ova* but of the *spermatozoa*). (6:80n., emphasis in original)

The specific seeds giving rise to radical evil, Kant argues here, must be passed on only through the male side. They therefore must be found in the sperm and not in the egg. Mary did not pass on the seeds she must have inherited from her father, and so Jesus was born without a propensity to evil. Radical evil is *gendered*; although both men and women have a propensity to evil, it comes essentially from male anatomy and is passed on through the male side.[26]

Commentators have generally passed over this footnote without comment; those who mention it tend to take it as being unserious. For example, Palmquist emphasizes the passage's conditional tense and wonders whether Kant means the whole thing as a *reductio*.[27] Similarly, McLaughlin interprets Kant as being "somewhat facetious" here, pointing to several pages of draft notes for the footnote that weigh possible explanations that are different from what appears in the text.[28] But the notes and drafts are more readily understood as evidence that Kant *was* serious, as he had carefully considered different options before settling on a solution. In his most recent book on religion, Wood

26. Differences in their theories of generation notwithstanding, this theory is remarkably similar to Augustine's. For Augustine's version of the argument that original sin is passed on through semen, see, e.g. Augustine (1966), *City of God, Volume iv: Books 12–15*, trans. Philip Levine, Loeb 414, Cambridge: Harvard University Press, 180–181 (this passage also includes a noteworthy description of the depravity of human nature as coming from a "diseased root" [*radice corrupta*]). For Augustine's application of this seminal transmission theory to the virgin birth, see e.g. book II, chapter 41 of Augustine, *Of the Grace of Christ and of Original Sin*, trans. Peter Holmes, accessed from the Logos Virtual Library.

27. Palmquist, Stephen (2014), *Comprehensive Commentary on Kant's Religion Within the Bounds of Bare Reason*, Oxford: Blackwell, 221n.21.

28. McLaughlin, Peter (2007), "Kant on Heredity and Adaptation," in *Heredity Produced: At the Crossroads of Biology, Politics, and Culture, 1500–1870*, eds. Staffan Müller-Wille and Hans-Jörg Rheinberger, Cambridge: MIT Press, 86. Kant's interesting notes are 23:105–108.

says of this passage only that "Kant resists any commitment to the letter of such Church doctrines as the virgin birth."[29] This comment is misleading, since it suggests that Kant rejects the theory; he in fact develops his own version of it. The footnote is important for my interpretation, however, because it states more directly than anywhere else that Kant sees the propensity to evil as part of human anatomy, even to the point of locating the corresponding "seeds" in sperm. It is hard to see how one could make any sense of this note or the draft material if one wanted to deny that the propensity to evil has its basis in human anatomy.

The second passage is in section IV of part one, which covers the genetic question of the origin of evil. Kant distinguishes between two senses of the term "origin" [*Ursprung*]: origin according to reason, and origin according to time. About the former Kant says that one must treat every evil action as if the perpetrator "had fallen into it directly from the state of innocence" (6:41), so that each moral transgression must be considered as its own spontaneous origin. He says this to preserve accountability: if an immoral act could be fully explained in terms of temporal antecedent causes, it would be a determined natural occurrence, for which the agent could not be held responsible. About the origin in time, however, Kant says the following:

> "In Adam we have all sinned" and still sin — except that a prior innate propensity to transgression is presupposed in us but not in the first human being, in whom rather innocence is presupposed with respect to time; hence his transgression is called a *fall into sin*, whereas ours is represented as resulting from a prior depravity of our nature. (6:42, emphasis in original)

Recall that "seeds" and "predispositions" are passed on through the generations at a certain level of development. In us the propensity to evil is "innate" [*angeboren*] (6:25; 6:29; 6:30), which is the same term

29. *Kant and Religion*, 149.

Kant uses to describe racial characteristics such as the skin color of Abyssinians (8:171). But in Adam this propensity was not innate, because he had no ancestors from whom he could have inherited it. His propensity must have been entirely undeveloped at first, untouched by any occasioning cause that would lead it to grow. His "fall into sin" gave this originally inactive propensity to evil the jolt it needed to germinate, just as the "humid heat" of Africa is supposed to have given the conditions for the original sprouting of the seeds associated with dark skin. Once the process started, its progress became irreversible. The only way a human being can avoid having a propensity to evil woven into the root of their nature is by being lucky enough to have God as their father (though not as their mother). The rest of us mortals inherit Adam's original sin — not as a moral fault for which we can be blamed, but as a corrupting propensity to evil. This propensity was dormant in him but active in us, and becomes more active with each generation. It is an advantage of my interpretation that I need not explain away these passages as ironic, facetious, or unserious.

6. The Propensity to Evil as a Racialized Characteristic

In the essays on race, Kant says that predispositions develop at the level of the species or the race, and not at the level of the individual. This is why he repeatedly insists in *Religion* that it is not individuals but the human *species* that is radically evil (6:21; 6:25; 6:29; 6:30). Because nature implanted predispositions in us as part of a wider system of ends, we ought to consider them from a broad teleological standpoint. Kant tells a story of historical progress, but it is slow and gradual: he does not expect to see much development in any one generation. This means that the extent of the development of one's seeds and predispositions depends proportionally less on the efforts one makes oneself over the course of life than on one's ancestral lineage. If one's descendants had nourished their propensity to evil during their own lives, feeding its growth by their wicked deeds, then one would inherit that characteristic as more fully developed already, making the temptation even stronger than it would be if one had been born into a different lineage. This view could never entail that a person be held morally responsible for acts done by another, but it does suggest that individuals could have a stronger or weaker propensity to evil based on the historic crimes of their ancestors. Similar considerations hold for the predisposition to the good. That we have a "predisposition to personality," which Kant distinguishes from personality itself (6:28), suggests that different individuals find it easier or more difficult to act out of respect for the moral law based on how far the predisposition has developed in them. As it unfolds, one would expect such people to do the right thing more often, but of course nothing follows about how morally praiseworthy their resulting actions are.

Read together with the natural history of humankind from the race essays, this account implies that these seeds and predispositions follow different developmental pathways within different groups of human beings, as long as those groups remain reproductively isolated from one another (as Kant believes human races generally have). Racial groups must therefore be at different stages of moral development on Kant's view, based on the relative strengths of the predispositions to good and propensities to evil they inherit from distinct ancestral lineages. This view is disturbing, but it is also in line with a specific form of racism one often finds in Kant, in which disfavored races are criticized for a supposed deficiency in the development of their innate predispositions.[30]

In fact, some parts of Kant's account of radical evil look like they would not apply to certain racial groups at all, given other things he says about them. For example, the reference to "vices of culture" (6:27), which are said to be grafted on to our predisposition to humanity, suggests that he was not thinking of indigenous Americans. Their "natural predispositions" [*Naturanlagen*], he believed, makes them "incapable of any culture" [*unfähig zu aller Cultur*] (8:176). In other words, because they lack the specific predisposition to develop culture in his view,

30. See, e.g., the comment in the Friedländer anthropology lectures, that a "savage Indian or Greenlander" has "the same seeds as a civilized human being, only they are not yet developed" (25:694).

they would also lack the corrupting propensity that turns culture rotten. That would suggest in turn that the account of radical evil is not supposed to apply to them in the way one might have expected.

To make sense of this comment, I claim that we ought to read Kant as telling two stories.[31] In the first, written for his predominantly white European audience, human beings are said to cultivate the seed of goodness, develop their predispositions, and advance through the history of freedom. This story includes the development of a propensity to evil that, somewhat counter-intuitively, actually helps humanity to develop its capacities. The predisposition to good and propensity to evil oppose each other in the individual, but from the historical perspective they work in tandem. Even radical evil was implanted in us by teleological nature as part of a wider plan that is ultimately for the best. As time proceeds, our seeds and predispositions continue to develop, and it becomes ever easier for people both to do good and to do evil, which grow together.

We learn what the second story, about nonwhites, might look like from an oft-discussed passage in the Herder review. In that review, Kant asks why Tahitians exist at all and wonders whether it would be "just as good" from the perspective of nature's ends if the island were populated instead with "happy sheep" (8:65).[32] We can understand Kant's reasoning by noting a passage from the paragraph of "Universal History," cited in section one. After defining "unsocial sociability," Kant considers what life would be like if human beings did not have this trait: "all talents would, in an arcadian pastoral life of perfect concord, contentment, and mutual love, remain eternally hidden in their seeds; human beings, as good-natured as the sheep they tended, would give their existence hardly any greater worth than that of their domesticated beasts" (8:21). For Kant, Tahitians remain in a Rousseauian state of contented indolence, and stand outside the universal story he tells about the progress of the human species as a whole within his biologized philosophy of history.[33] The Tahitians' lack of interaction with so-called "cultured" people keeps them isolated socially but also sexually; their seeds and predispositions can therefore not develop. As such they do not grow their talents, but they also avoid the many new problems accompanying social development, which is why Kant calls them "happy" even as he dehumanizes them by comparing them to animals. As descendants of Adam, Tahitians must have a propensity to evil, but from Kant's description it appears that this propensity has hardly developed at all. Kant's theory of radical evil may well be universal in its form, but as Mills has argued about other elements of Kant's moral philosophy, it does not follow that the theory was intended to apply to all human beings in the same way. Radical evil is racialized.

7. Conclusion

In this paper, I have developed a new anthropological interpretation of Kant's theory of radical evil. This interpretation aims to take Kant's organic language seriously. While there have been other "anthropological interpretations," none have explored the link with Kant's race theory, which gives *Religion* its core concepts and even section headings. My interpretation reaffirms something that dismayed some of the book's early readers: that it is indeed a new version of the doctrine of original sin. Like Augustine, Kant develops an intricate quasi-biological theory of original sin's transmission via semen. The theory exempts Jesus and Adam, although unlike Augustine, Kant takes human beings to inherit a mere *propensity* to evil, rather than moral guilt proper. I have also argued that we should understand propensities and predispositions as intrinsically developmental concepts which exist to a different degree in each individual. To have a stronger propensity to evil is to have something in oneself making temptation even more

31. I thank an anonymous reviewer for the idea of putting it this way.

32. On this passage, see Bernasconi, Robert (2005), "Why Do the Happy Inhabitants of Tahiti Bother to Exist at All?", in *Genocide and Human Rights: A Philosophical Guide* ed. John K. Roth, New York: Palgrave Macmillan, 139–148, and Harfouch, John (2018), *Another Mind-Body Problem: A History of Racial Non-Being*, Albany: SUNY Press, 141–146.

33. On the racialization of "laziness" in Kant, see Lu-Adler, Huaping (2022), "Kant on Lazy Savagery, Racialised," *Journal of the History of Philosophy* 60(2), 253–275.

difficult to resist; to have a stronger predisposition to the good is to have something in oneself making it easier to do what one ought. By taking seriously the link to Kant's account of heredity implied by his use of these concepts, I have argued that this development happens slowly, which means that its current status is ultimately determined more by one's ancestral lineage than by one's own efforts.

The reference to race, however, also reveals a darker side to Kant's theory, which is that the theory cannot be assumed to apply in the same way to all human groups. Although *Religion*'s main claims are universal in their form, their anthropological basis means that they belong to Kant's natural history of the human species, which is where one finds many of his most racist ideas. Previous anthropological interpretations have connected radical evil to Kant's philosophy of history, but that philosophy of history faces serious problems that have not been part of the discussion. My interpretation does not aim to exonerate Kant; my goal has been to develop an interpretation that best fits with our understanding of the rest of his work. This includes his racism, which often takes the form of disparaging comments about groups who are said to be lacking in the development of their seeds and natural predispositions.

I would like to end by calling for *Religion* to play a more prominent role in debates about Kant and race than it has thus far. Its extensive reference to seeds and predispositions shows that Kant did not stop using the main concepts of his race theory in the 1790s, as some have suggested.[34] More work remains to be done to say exactly what this point means for his views on generation and heredity after the third *Critique*, but the use of the terms in *Religion* is at least *prima facie* evidence that he did not simply abandon them. It also shows that even into his late period, Kant was still finding new uses for what turned out to be a surprisingly malleable and adaptable theory.[35] Indeed, *Religion* introduces these ideas into a context in which they had never before appeared in Kant's work: his moral philosophy. Although Kant makes statements that have been characterized as racist in many of his works, he does not systematically appeal to the concepts of his race theory in any major writings on moral philosophy in the 1770s and 1780s. But he does in the 1790s.[36] *Religion* includes what may be the closest link between Kant's moral philosophy and his race theory, though it has played little role in the debate so far.

34. See Kleingeld, Pauline (2007), "Kant's Second Thoughts on Race," *The Philosophical Quarterly* 57(229), 591–2, esp. n.40 and Sloan, "Preforming the Categories," 253.

35. On which, see Sandford, Stella (2018), "Kant, Race, and Natural History," *Philosophy & Social Criticism* 44(9), 950–977.

36. The focus of this paper has been *Religion*, but it is worth mentioning that while "seeds" disappear from Kant's work after 1793, "predisposition" remains part of his lexicon even in other late texts on moral philosophy, such as the *Metaphysics of Morals*. Though not as prominent as in *Religion*, Kant uses the concept some 18 times over the course of the book.