

Judging The Inner Judge: On The Genesis Of Kant's Theodicy Essay

Pablo Muchnik
Emerson College



<https://dx.doi.org/10.5209/kant.104653>

Accepted: 22-12-2025

ENG Abstract: This paper interprets Kant's 1791 essay, "On the Miscarriage of All Philosophical Trials in Theodicy," as an exercise in philosophical self-examination. Instead of attributing the rejection of doctrinal theodicy to the maturation of his moral theory or the impact of external influences (as other interpreters have done), I read *Miscarriage* as Kant's belated attempt to restrain his own rationalist tendencies even after writing the *Critique of Pure Reason*. Those tendencies emerge throughout the 1780s in his lectures on religion and writings on history. Kant repudiates them in *Miscarriage*, recasting the figure of Job as a model of "authentic theodicy," a man who renounces *knowledge* claims about God in favor of his *moral* integrity. This move upends Kant's earlier theodicean strategies and curbs the temptation that led him to embrace them. So construed, Kant's 1791 essay can be read as a tribunal in which he judges his prior rationalist self – a philosophical autobiography of sorts.

Keywords: Kant, theodicy, Job, critical philosophy, speculative reason, transcendental illusion

Palabras clave: Kant, teodicea, Job, filosofía crítica, razón especulativa, ilusión transcendental

Summary: I; II; III; IV; V; VI; Bibliography.

How to cite: Muchnik, P. (2026). Judging The Inner Judge: On The Genesis Of Kant's Theodicy Essay. *Con-Textos Kantianos. International Journal of Philosophy*, 23, 19-29. <https://dx.doi.org/10.5209/kant.104653>

Those who engage in doctrinal theodicy, Kant tells us at the beginning of "On the Miscarriage of All Philosophical Trials in Theodicy" (1791), misconstrue what they are doing.¹ Although they believe to be "[defending] the highest wisdom of the creator against the charge which reason brings against it for whatever is counterpurposive in the world" (MpVT 8: 255), they are in fact advancing a specious cause of their own making. For their project, Kant explains, results from a fatal combination of self-conceit and lack of self-knowledge: the references they make to God's wisdom are the product of their "presumptuous reason failing to recognize its own limitations" (*ibid.*). This disregard for limits, we learn as the essay proceeds, is not a simple cognitive mistake. It is also, and more importantly, expression of a dubious moral character, a sign of the mendacity that the friends of Job embody and which Kant considers "the principal affliction of human nature" (MpVT 8: 267).

The charge of epistemic overreach is precisely what one would expect from Kant's Copernican turn in the first *Critique*. Confining knowledge to the bounds of possible experience gives a mortal blow to the rationalist, Leibnizian-style projects Kant undertook throughout the precritical period.² Yet, it would be both misleading and simple-minded to think that Kant's theodicean attempts could be cleanly partitioned into a dogmatic pre-critical stage and a more circumspect view of the speculative power of reason in the postcritical period. Indeed, as various interpreters have indicated, Kant actively endorsed, long after publishing the *Critique of Pure Reason* in 1781, many of the same views he came to reject in the Theodicy essay.³

This inconsistency is most glaring in his *Lectures on the Philosophical Doctrine of Religion* (1783-84), but is also present in his writings on history, particularly "Idea for Universal History" (1784) and "Conjectural

¹ "Doctrinal theodicy" is the term of art Kant uses in *Miscarriage* to refer to any attempt that resorts to speculative reason to vindicate God's wisdom (MpVT 8: 264). It designates the style of theodicy Leibniz made famous in *Essais de Théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal* (1710), and which became dominant in the rationalist tradition.

² George Huxford reconstructs the evolution of Kant's thoughts on these matters. See George Huxford, *Kant and Theodicy: A Search for an Answer to the Problem of Evil* (Lexington Books: Lanham, Maryland, 2020).

³ See Christoph Schulte, "Zweckwideriges in der Erfahrung: Zur Genese des Mißlingens aller philosophischen Versuchen in der Theodicee bei Kant," *Kant Studien* 82 (1991): 371-3, 382. The same chronological discrepancy has been noted, among others, by Georg Cavallar (see "Kants Weg von der Theodizee zur Anthropodizee und retour: Verspätete Kritik an Odo Maquard," *Kant Studien* 84 (1993), 90-102) and Johannes Brachtendorf (see "Kants Theodizee Aufsatz – Die Bedingungen des Gelingens philosophischer Theodizee," *Kant Studien* (2002) 93, 90-102).

Beginning of Human History" (1786).⁴ In these published texts, Kant rehearses the same theodicean strategies his students learned in his classroom. It would be tempting to shift the blame for this inner tension to Baumgarten's *Metaphysica*, the manual Kant used in much of his *Lectures*.⁵ But that would be a mistake. As I will indicate below (section IV), it is Kant himself, under his own pen and at the height of his intellectual powers, who refuses to abandon some of the speculative habits he criticized in his rationalist predecessors.

How are we to interpret the coexistence of critical and transcendent tendencies within Kant's mature philosophizing? What explains Kant's stubborn adherence to doctrinal theodicies despite his commitment to the principles of Transcendental Idealism? Why does he condemn as morally dubious an intellectual project he previously embraced? And why did he decide to abandon it? The goal of this paper is to propose a reading of the 1791 essay that answers these questions.



Let me begin with the last query. The problem of why Kant eventually gave up on doctrinal theodicy has received two kinds of answers in the literature. The first appeals to external influences. This is the view advanced, for example, by Sam Duncan in "Moral Evil, Freedom, and the Goodness of God: Why Kant Abandoned Theodicy."⁶ As Duncan sees it, it is the work of Carl Christian Eberhard Schmid that forced Kant to confront a serious flaw in his prior understanding of evil, namely, his claim that immorality is caused by human limitations. This line of thought, Schmid indicated in *Versuch einer Moralphilosophie* (1790), was deeply problematic: wrongdoing so construed is devoid of any moral import, since the failure to act morally is explained by an appeal to "factors outside of us and outside our control."⁷ Given Kant's identification of freedom and morality in the *Groundwork* (1785), this account became untenable: since the imperfections attached to our finitude depend on God's arrangement of creation, we cannot be blamed for the immoral acts we perform.⁸ Any attempt to refer immorality back to the divine architecture of the universe, Kant eventually realized, may serve to vindicate God but at the price of denying human agency in wrongdoing.

The second type of answer takes the opposite route and traces the genesis of the *Theodicy* essay back to Kant's internal philosophical development. The account by Robert Gressis is emblematic of this approach. In "Kant's Theodicy and its Role in the Development of Radical Evil," Gressis tackles what appears to be a major contradiction in Kantian ethics: Kant's shifting views on inclinations.⁹ While in the *Groundwork* (1785) and the second *Critique* (1788) Kant considered inclinations to be both undesirable and blameworthy, in *Religion* (1793) he claims that inclinations are good in themselves and that any attempt to extirpate them is reprehensible and futile. There is, for Gressis, no inconsistency between these claims: Kant made them both –what changed between these two periods was his understanding of immorality. While in the 1780s Kant saw inclinations as tempters to be vanquished (and if possible, eradicated), by *Religion* (1793) he had developed the doctrine of radical evil, according to which evil results from our collusion with temptation. So conceived, evil is not an enemy that assails us from without but an enemy within, a self-imposed propensity which corrupts our moral judgment and welcomes the attack. Theodicean reasons, Gressis argues, underpin Kant's change of mind: appealing to the "essence of things" to explain immorality, as Kant did in 1780s, could no longer be an acceptable philosophical move, but became, instead, another subterfuge of radical evil, a clever attempt to evade our responsibility for wrongdoing.

I find much to commend in these interpretations. Despite following different interpretative paths, Gressis and Duncan call our attention to a key issue: Kant was forced to wake up from his "theological slumber" when he realized that, in the process of vindicating God's moral wisdom, doctrinal theodicies end up explaining evil away.¹⁰

It should be clear, however, that no matter their merits, these readings are sorely lacking. Concerned as they are with reconstructing the reasons of Kant's evolving thoughts on immorality, they lose sight of what is perhaps the most perplexing aspect of the 1791 essay. Namely, they ignore the fact that, for a good part of a decade *after* publishing the first *Critique*, Kant clung to a philosophical project that egregiously violated the cognitive constraints he had set for human reason –a project which he knew (or should have known) was

⁴ One may also want to include *Perpetual Peace* (1795) on this list, for in this text Kant interprets war as an engine of institutional and moral progress, i.e., as an unwitting instrument for the good. Thus, both the destructiveness of armed conflict and the vices that accompany it are viewed as part of a providential/divine plan. Seen this way, Kant's theodicean impulses do not die with the *Theodicy* essay but haunt him well into the 1790s. What changes at this later stage of Kant's career is the status of the judgments used to vindicate God's wisdom: they are no longer constitutive, as they were in the prior decade, but consistently reflective judgments.

⁵ Kant, like all professors at the time, was required to adopt in his courses manuals that had been sanctioned by the authorities. He used Baumgarten's *Metaphysica* (Part IV, "Theologia Naturalis", §§800-1000) to organize the bulk of his *Lectures on the Philosophical Doctrine of Religion*, except for the Introduction, which drew primarily from Eberhard's *Vorbereitung zur natürlichen Theologie zum Gebrauch akademischer Vorlesungen*. Kant's *Lectures* were eventually published by Ludwig Pölitz in 1817, long after his death.

⁶ Sam Duncan, "Moral Evil, Freedom, and the Goodness of God: Why Kant Abandoned Theodicy," *British Journal for the History of Philosophy* 20 (5), 2012: 973-991.

⁷ Duncan 2012, 982.

⁸ Duncan 2012, 988. One can hear echoes in Duncan's historical reconstruction of the debate between Henry Allison and Gerold Prauss regarding the role of Reinhold's *Letters* in the development of Kant's conception of freedom. For Duncan, however, it is Schmid (not Reinhold) who forced Kant to revise the assumptions of the so-called analyticity thesis. See Henry Allison, *Kant's Theory of Freedom* (Cambridge, 1990), pp. 133 ff.

⁹ Robert Gressis, "Kant's Theodicy and its Role in the Development of Radical Evil," *Archiv für Geschichte der Philosophie* 100 (1), 2018, 46.

¹⁰ See MpVT 8:259.

bankrupt from the start. Even if these interpretations got things right and explained why Kant ultimately abandoned this undertaking, they leave unanswered the logically prior question of why he adopted it to begin with.

To avoid this shortcoming, we need to develop an interpretation that retraces Kant's steps and places us back at the point where his critical commitments begin to wobble, i.e., the moment when the temptation to justify why God put up with evil in the world emerges with irresistible force. For we do not simply want to know why Kant *renounced* doctrinal theodicy. Equally important, and arguably more intriguing, is to know why he *embraced* it in the first place, disregarding the advice of his better philosophical angels. We need, in short, an account that sheds light on what current readings brush aside, that is, an interpretation which helps us understand why Kant found the theodicean appeal so enticing, surrendered to its spell, and remained for so long oblivious to having transgressed any epistemic boundary.

-II-

In "Miscarriage," Kant provides a clue to interpret his curious behavior:

This [i.e., doctrinal theodicy] is indeed not the best of causes, yet one can be condoned insofar as (aside from that self-conceit ["of our presumptuous reason failing to recognize its limitations"]) the human being is justified, as rational, in testing all claims, all doctrines which impose respect upon him, before he submits himself to them, so that this respect may be sincere and not feigned. (MpVT 8: 255)

Kant's account, however, is at once disingenuous and revealing. It is disingenuous, for it makes doctrinal theodicies appear, in the very act of violation, as byproducts of the same critical spirit that gave rise to his *Transcendental Philosophy*. He described that spirit in the A Preface to the first *Critique*:

Our age is the genuine age of **criticism**, to which everything must submit. **Religion** through its **holiness** and **legislation** through its **majesty** commonly seek to exempt themselves from it. But in this way they excite a just suspicion against themselves, and cannot lay claim to that unfeigned respect that reason grants only to that which has been able to withstand its free and public examination. (KrV A XI n.)

If, as we indicated above, Kant came to see doctrinal theodicies as sly intellectual maneuvers to disavow our immorality, the rational "testing of claims" he adduces in "Miscarriage" as a justification for his prior conduct is an attempt to conceal what, in retrospect, must have been a cause of philosophical embarrassment for him. After all, it was not the self-conceit of human reason as such, but the self-conceit of *his own* reason, which, just a few years before, enticed him to take flight into the supersensible and think that the problem of evil could be solved by *understanding* God's intentions. The entreaty to put aside the arrogance that motivated his own project, and thus manufacture the conditions favorable to condone it, is precisely what he was in no position to do without prevaricating. If this reading is correct, it should come as no surprise that neither in the *Theodicy* essay, nor, to my knowledge, anywhere else in the corpus, did Kant acknowledge having committed a philosophical transgression in those earlier years.

But there is more than evasion here. Kant's identification of what induces a philosopher to embark in theodicy with the activity of examining the validity of claims that command our respect, i.e., the venerable activity he calls "critique," is also revealing. It amounts to a slip of the tongue, i.e., an unintentional revelation of what motivated Kant to write this essay and how he wished *Miscarriage* to be read. For, if we heed to what these passages are intimating, we'll see that Kant is secretly expressing a desire to deploy the tools of critical philosophy upon himself. Such deployment is necessary, Kant eventually realized, to overcome the rationalist presumptions that not long ago led him to engage in doctrinal theodicies. Seen this way, what Kant is telling us, though in an oblique and distorted fashion, is that he wished *Miscarriage* to be read as a therapeutic exercise in self-criticism, as a belated attempt to restrain his own rationalism and submit his philosophy of religion to the harsh discipline of critique.

Such a discipline is necessary, I argue toward the end of this paper (section V), because Kant came to see doctrinal theodicies as akin to transcendental illusions, i.e., as a variant of the kind of error Kant's rationalist predecessors typically made. To resist the allure of those illusions, Kant believes, one needs to erect "a system of cautions and self-examination out of the nature of reason and the objects of its pure use, before which no false sophistical illusion can stand up but must rather immediately betray itself" (KrV A711/B739).

The target and scope of this system of safeguards is, of course, different in *Miscarriage* than in the first *Critique*. While in the *Critique* it was theoretical reason which came under surveillance and self-examination, in the *Theodicy* essay the target is more personal, the subject matter more narrowly conceived. For here Kant wants to bring the tools of criticism to bear upon himself. His goal is not to put metaphysics into the secure path of science, but to test the validity of the assumptions of his own philosophy of religion. Kant's project in *Miscarriage*, to put it bluntly, is deliberately reflexive and self-referential. He is engaged in a philosophical autobiography of sorts, a self-examination whose point is not to scrutinize the workings of human rationality in the abstract, but the soundness of his earlier theological convictions. It is not reason in general that is being disciplined here – it is Kant himself.

-III-

Although self-discipline is always painful, it also brings important rewards for those willing to undertake it. As we learned in the first *Critique*, it yields a Socratic form of wisdom, a "negative wisdom" which is very much like Job's – a figure whom Kant conceives in the likeness of the Greek hero, since both characters have

“insight into the necessary limitation of what we may presume with respect to that which is too high for us” (MpVT 8: 263). Like Socrates, Job mastered the main lesson of “negative instruction,” and thus acquired the wisdom to *know that he does not know*. This is precisely the lesson the late Kant seeks to impart to himself in *Miscarriage*. He was now ready to receive it for by the 1790s he finally realized that knowledge of one’s ignorance was necessary “to bring this trial [of God’s wisdom] to an end *once and for all*” (MpVT 8:263). What drives this undertaking, I would like to suggest, is the need to vanquish the temptations that led him to adopt his prior theodicean views, an adoption which in hindsight must have appeared to him as an affront to his own critical commitments.

Kant was able to make philosophical amends, we may conjecture, only after the rationalist part of himself, i.e., the part that had tempted him to overstep the limits of human cognition and engage in doctrinal theodicy, accepted that knowledge of one’s ignorance is not a form of failure, a mere humiliation to the majesty of reason. This is something the critical Kant had ascertained from the very beginning, but which took him a long time to take to heart, for it entailed the wholesale rejection of an ingrained assumption in the philosophical tradition, namely, the assumption that metaphysics is first philosophy. Fundamental assumptions like this are hard to abandon; they refuse to go away, even after we recognize the price we pay to hold them. Indeed, it is only with doctrine of the primacy of practical reason, articulated in the second *Critique* towards the end of the 1780s, that Kant finally turns that metaphysical tradition on its head. For, among other things, this doctrine compensates cognitive renunciation with overriding practical gains: it teaches us that the purpose of philosophy is not to make deep theoretical pronouncements about ultimate reality, but to help us lead a virtuous life. And what is true for Transcendental Idealism writ large also holds for Job in matters of religion, where the desire to *know* God through speculative reason must be deflected into a strictly moral *faith* in Him.

Forfeiting one’s speculative pretensions, the late Kant understood, is not a loss. As the case of Job is meant to illustrate, it is a sacrifice far outweighed by the moral benefits it entails. For, in a world plagued by the threat of the counterpurposive, Kant takes Job to teach us that only an unflinching faith in God will allow us to preserve our moral commitments despite great misfortune. It is this faith that sets Job apart from Spinoza, the righteous atheist Kant described in the third *Critique* (1790), who, like Job “actively honors” the moral law, but unlike him believes that “there is no God” (KU 5:452).¹¹ Surrounded by “deceit, violence, and envy,” in a world where we can expect to be visited with “the evils of deprivation, disease, and untimely death” regardless of our moral worth, Kant believes that the righteous atheist will break under pressure, yield to despair, and “give up as impossible” what the moral law obligates him to pursue (KU 5: 453). In Kant’s eyes, Spinoza’s atheism, which is as dogmatic as the theism of rationalist theologians, will inevitably weaken the respect he feels for morality –very much as it does with Job’s friends. What binds these disparate characters together, Kant reckoned, is that they all profess to have a certainty beyond their reach, and hence are trapped in an epistemic game they are not equipped to play, for whether “there is no God” or “there is a God that rewards our obedience,” are assertions that say more than their authors know.

Job dodges this bogus epistemic game altogether, for his faith has long forsworn any pretense of knowledge and accepted the cognitive inaccessibility of God’s will. This attitude underlies the “system of *unconditional divine decision*,” which Job espouses by saying, “‘He has decided’ ... ‘He does as he wills’” (MpVT 8:265). Job’s words are not directed to God but to himself. They do not mean to signify God’s omnipotence or the mysterious ways of His justice. They are meant, rather, as an indication of his moral character and a reminder of the impotence of reason to understand the laws that structure the world.¹² Uttering them, Job gives proof of the drastic reorientation his character has undergone. For in him, the drive to engage in transcendent speculation, i.e., the drive that still holds sway both on Spinoza and Job’s friends, has been curbed and redirected to immanent practical purposes.

In Kant’s mind, this reorientation of one’s rational drives gives rise to a conviction that also deserves “the name of ‘theodicy,’” for it renders moot “all objections to divine wisdom” (MpVT 8: 264). Kant calls this type of theodicy “authentic” to distinguish it from the “doctrinal” approach of its speculative counterpart (MpVT 8: 264). Here, “less depends on subtle reasoning (*aufs Vernünfteln*) than on sincerity in taking notice of the impotence of our reason, and on honesty in not distorting our thoughts in what we say, however pious our intention” (MpVT 8: 267). For, turning away from the falsifying effects of “*rationalizing (vernunftelnden)* reason,”¹³ Job lets his “*efficacious* practical reason” have the last word. It is the moral quality of his life, not his religious assertions, rectitude and not orthodoxy, which serve as “the unmediated definition and voice of God through which he gives meaning to the letter of his creation” (MpVT 8:264). For Spinoza and Job’s friends, by contrast, God remains a transcendent object of speculation, a morally inert abstraction.

The Kantian motto, “I had to deny knowledge in order to make room for faith” (KrV BXXX), underpins and makes possible Job’s attitude toward God. The denial of knowledge does not simply contain a cry for

¹¹ In addition to denying God’s existence, the righteous atheist believes that “there is no future life” (KU 5: 452). He thereby rejects the two assumptions supporting the postulates Kant uses to build the doctrine of the highest good, and on its basis, his philosophy of religion. Since Kant does not ascribe to Job the belief in immortality, I will limit my discussion to matters of theology, where the contrast with Spinoza is most conspicuous.

¹² I here disagree with Charles Larrimore, who takes Job’s position to indicate “an unabashed voluntarism.” Larrimore considers Job’s friends, on the other hand, to be “antivoluntarists,” for they “argue from ‘that system which explains all ills in the world from God’s justice,’ while Job ‘declares himself for the system of unconditional divine decision.’” Putting things this way, I believe, completely misconstrues the role Job is playing in Kant’s interpretation. See Charles Larrimore, “Autonomy and the Invention of Theodicy,” in *New Essays on the History of Autonomy: A Collection Honoring J.B. Schneewind*, Natalie Brender and Larry Krasnoff (eds.), Cambridge University Press, 2004, p. 81.

¹³ The English translation renders “*vernunfteln*” as “ratiocinating.” I find “rationalizing” a better turn of phrase.

epistemic humility, but also the key to remaining virtuous in a world that is indifferent to our moral striving. Thus, Kant finds reflected in Job's uprightness an answer to the question, "If I do what I should, what may I then hope for?" (KrV A805/B833) –albeit, in this case, an answer that arises from confronting the counter-purposive, and not, as it did in the prior *Critiques*, by trying to integrate the disparate demands of happiness and duty.¹⁴ In Job's response to the inscrutable reasons behind his misfortune, Kant sees an alternative way to justify his claim that "morality inevitably leads to religion" (RGV 6:6).¹⁵ But the pathway here is not through the *summum bonum*, as is customary in the Kantian corpus, but through its conceptual counterpart, the *summum malum*, which forces us to make an existential choice that cannot be justified on epistemic grounds: the choice between religion and despair.¹⁶

To reach a decision, Kant surmises, Job must have found a pathway to avoid both the cognitive excesses of Spinoza's atheism and the immorality of his friends. For, while in the grip of great affliction, the righteous atheist exclaims, "Since there is no God, my torments are unbearable," Job's friends, smug in their good fortune (MpVT 8:265), arrogantly assert, "Since God exists, there is no innocent suffering."¹⁷ Each character thinks, Kant tells us, "according to [their] station (*Lage*)" (MpVT 8:265), conceiving of God in light of their particular condition. And so does Job, who avoids both the dogmatism and the bad faith behind such claims. He remains, instead, wisely silent about the causes of his ordeal, refusing the temptation to either blame himself (as his friends insist he should) or surrender to despair (as the righteous atheist does). Moral self-abasement and philosophical self-conceit are twin vices for Kant: they prevent us from experiencing the indignation a good person must feel towards a suffering she does not deserve, but to which she is unwilling to capitulate.

Now, if we step back from the text and reflect on the principles of its composition, we'll realize that what makes Kant's use of archetypal figures in *Miscarriage* so effective is that critical discipline is a reflexive activity. Seen that way, Job and his friends can be interpreted as symbols that allow Kant to talk about himself, to name and confront an inner tension that would have been otherwise inarticulable for him. Through these figures, Kant brings a latent psychic conflict to open philosophical expression. This transfiguration has profound therapeutic effects. By its means, Kant manages to integrate aspects of himself that have been warring for a decade, namely, his critical and his rationalist philosophical commitments –personified, respectively, by Job and his friends.

The path of indirection, however, is not limited to the autobiographical. It also shapes the formal structure of this essay. It explains, for example, why Kant resorted to the metaphor of a "juridical process ... before the tribunal of reason" (MpVT 8: 255) as the guiding trope in *Miscarriage*. This metaphor not only plays a major architectonic role in the text, structuring the first part as a series of exhibits submitted to the court, and the second as corollaries of the verdict. In addition, and equally significant for the interpretation we are proposing, the imagery of a tribunal is well suited to capture the internal split Kant must have felt in the process of self-examination. For if Kant, as we are claiming, uses *Miscarriage* as an occasion for "negative instruction," then the figures of the "plaintiff," the "advocate," and the "judge" should not be construed as referring to three different persons. They must be seen, instead, as various aspects of Kant himself, who simultaneously is the object and the subject of critical scrutiny. To the extent that Kant's primary goal in *Miscarriage* is to discipline the urge that led him to engage in doctrinal theodicy, we must conclude that it is his earlier rationalist self that is here being judged by its outgrown critical incarnation.

Thus, despite appearances to the contrary, the officers of the court should not be seen as separate agencies. In the tribunal of reason, they are masks Kant wears throughout the trial, literary devices that serve him as props to discipline himself. After all, Kant is no stranger to disguises: he employs them with "Job," his critical alter ego, and with "Job's friends," who voice theological views Kant had been prone to endorse in the recent past and now actively repudiates. Without exception, all these characters are fleeting shadows of a single, multifaceted subject, projections of a theodicean drama that is being played within Kant himself.

-IV-

To bring my argument to completion, I need to make good on two claims I have not yet had a chance to substantiate:

1. That *Miscarriage* presents a version of the same theodicean views Kant had embraced in the 1780s, and therefore that it is Kant's own rationalism which stands here on trial.
2. That the appeal of those views persisted, despite their spurious epistemic and moral implications, because doctrinal theodicies behave like transcendental illusions.

¹⁴ Lawrence Pasternack makes a similar point in "Kant's Moral Argument and the Problem of Evil: Authentic Theodicy and the Sincerity of Faith," in *Kant on Proofs on the Existence of God* (Ina Goy ed.). (Walter DeGruyter, 2022).

¹⁵ Kant makes this claim also in KpV 5:129.

¹⁶ As I see it, the path to Kantian religion runs through two parallel tracks: one track is driven by the question of what *should* happen to the virtuous (but does not), the other by the question of what *shouldn't* happen (but does). The former is the path of *hope*, the latter, of moral *indignation*. These paths are connected, respectively, to the *summum bonum* and the *summum malum*. Although both give rise to moral faith, one clings to the notion of divine justice while the other accepts the inscrutability of God's will; one makes us feel at home in the universe, the other captures our alienation from it.

¹⁷ For similar formulations, see Kenneth Seeskin, "Job and the Problem of Evil," *Philosophy and Literature*, Volume 11. 2 (October 1987): 230-40.

The combined effect of these claims is designed to shed light on what other interpretations ignore, namely, the troublesome coexistence of transcendent and critical tendencies within Kant's mature philosophizing. In this section, I will provide textual evidence in support of this view; in the next, I will discuss the etiology of transcendental illusions to explain why the conflict remained for so long undetected.

At the start of *Miscarriage*, Kant tells us that doctrinal theodicies must resort to one of these three argumentative strategies:

For this vindication [of the highest wisdom of the creator] it is required that the would-be advocate of God prove *either* that whatever in the world we judge counterpurposive is not so; *or* if there is any such thing, it must be judged not at all as an intended effect but as the unavoidable consequence of the nature of things; *or*, finally, that it must at least be considered not as the intended effect of the creator of all things but, rather, merely of those beings in the world to whom something can be inputted, i.e., of human beings (i.e., higher spiritual beings as well, good or evil, as the case may be). (MpVT 8:255)

These strategies provide the framework for the first part of Kant's essay: they give rise to the nine theodicean theses that result from systematically applying these lines of argument to the three kinds of counterpurposiveness (i.e., evil, pain, and the disproportion between them) that threaten to undermine God's moral attributes (i.e., his holiness as lawgiver, his goodness as ruler, and his justice as judge).

Although there is no sign of the nine theses in the *Lectures on the Philosophical Doctrine of Religion*, the text is structured according to the same organizational plan.¹⁸ As in *Miscarriage* (MpVT 8: 257 n.), Kant's first move is to dismiss physicotheology in favor of "a moral concept of God," which alone provides a determinate conception of the divinity (V-Phil-Th/Pöhlitz 28: 1073). Also as in "Miscarriage" (MpVT 8:257), Kant holds that God "must possess without limit the moral perfections of *holiness, benevolence, and justice*," which together "constitute the entire moral concept of God" (V-Phil-Th/Pöhlitz 28:1073).¹⁹ Furthermore, in both texts he insists that the lexical order between these attributes must be preserved at all costs: "We must think of the holiness of the laws as *first*, even though our interest commonly beguiles us into placing God's benevolence above it" (V-Phil-Th/Pöhlitz 28:1074). If we were to alter this ranking, divine justice would be seriously compromised, for a benevolence "without limit" would pervert the distribution "of happiness according to the proportion of worthiness in the subject" (*ibid.*).

At this juncture, however, Kant's attitude towards doctrinal theodicy drastically changes. With an assurance that the late Kant would have found astounding, he proceeds in the *Lectures* to dismiss as ungrounded the "objections whose strength have driven many human beings crazy and plunged them into despair" (V-Phil-Th/Pöhlitz 28:1076):

The first objection is against God's *holiness*. If God is holy and hates evil, then whence comes this evil, which is an object of aversion to all rational beings and is the ground of all intellectual aversion?

The second objection is against God's *benevolence*. If God is benevolent and wills that human beings be happy, then whence comes all the ill in the world, which is an object of aversion to everyone who meets with it and constitutes the ground of physical aversion?

The third objection is against God's *justice*. If God is just, then whence comes the unequal apportionment of good and evil in the world, standing in no community with morality? (V-Phil-Th/Pöhlitz 28:1076-7).

To respond to these charges, Kant adopts an approach that combines the first two justificatory strategies he mentions in *Miscarriage*. In a sweeping tale of moral progress, he sets himself to demonstrate "*either that whatever in the world we judge counterpurposive is not so, or if there is any such thing, it must be judged not at all as an intended effect but as the unavoidable consequence of the nature of things*" (MpVT 8: 255). Kant's story goes like this: "nothing can arise without its first predisposition being made by its creator" (V-Phil-Th/Pöhlitz 8: 1077). This does not mean, however, that "God himself placed a predisposition to evil in human nature" (*ibid.*). On the contrary, God shows his wisdom by having made the human being the only creature "who has to work for his perfection and for the goodness of his character, producing them from within himself" (*ibid.*). This is why "He created the human being free ... but gave him also animal instincts ... to be moderated and overcome through the education of his understanding" (*ibid.*). Although "perfect both in his nature and regarding his predispositions," with respect to "their education [human beings remain] still uncultivated" (*ibid.*). Their task is to leave behind this benighted state and "free [themselves] from [their] instincts" (*ibid.*). Given their lopsided nature, however, they are ill-equipped for the challenge, so that there await them "[o]nly false steps and foolishness" (*ibid.*).

"This way of representing things," Kant maintains, "agrees completely with the Mosaic story," according to which the human being lives in paradise, "undisturbed by his instincts, until finally he feels his humanity, and in order to prove his freedom he *falls*" (V-Phil-Th/Pöhlitz 28: 1078). At this point, "he no longer is an animal, but has *become* an animal" (*ibid.*). In Kant's account, the need for education and the misuse of freedom are inextricable: "with each new step he takes some new false steps, and in this way he approaches ever nearer to the idea of perfection in a rational being" (*ibid.*). "[E]vil in the world," it follows, "can be regarded as *incompleteness in the development of the germ toward the good*," for it "has *no special germ*; it is *mere privation*

¹⁸ Kant's discussion of theodicy takes place in the Second Part of the *Lectures* (under the title "Moral Theology"), particularly in the first section, "On the Moral attributes of God" (V-Phil-Th/Pöhlitz 28:1071-82). As in *Miscarriage*, the second section deals with matters of faith and is called "Of the nature and certainty of moral faith" (V-Phil-Th/Pöhlitz 28:1082-86).

¹⁹ Throughout this discussion, I preserved the italics in the original.

and consists only in the *limitation of the good*" (V-Phil-Th/Pöhlitz 28: 1078). Only the good "*has a germ* [that is] *self-sufficient*": "[a] *special germ toward evil cannot be thought*, but rather *the first development of our reason toward the good is the origin of evil*" (*ibid.*). This is so, because God "*wills that evil be removed through the progress toward good*": once "*the human being has finally developed himself completely, evil will cease of itself*" (V-Phil-Th/Pöhlitz 28: 1079). Kant concludes his defense: "*This justifies God's holiness, because by following this path the whole species of the human race will finally attain to perfection. But if we ask where the evil in the individual human being comes from, the answer is that it exists on account of the limits necessary to every creature*" (V-Phil-Th/Pöhlitz 28: 1079).

Something similar could be said about the two other forms of counterpurposiveness, pain and the unjust distribution of rewards and punishments. But this brief reconstruction suffices to show that Kant's theodicean views in the *Lectures* can be attributed neither to Baumgarten nor to the clumsy note-taking skills of Kant's students. It is Kant himself who espouses them, as he makes clear in his writings on history. In "Conjectural Beginning" (1786), for example, Kant hones on the same "Mosaic story" he mentions in the *Lectures*:

The result of this presentation of the first history of human beings is that the departure of the human being from the paradise which reason represents to him as the first abode of his species was nothing other than the transition from the crudity of a merely animal creature into humanity, from the go-cart of instinct to the guidance of reason—in a word, from the guardianship of nature into the condition of freedom. (MAM 8: 115)

Adding a few lines later:

Nevertheless, this course, which for the species is *progress* from worse toward better, is not the same for the individual. Before reason awoke, there was neither command nor prohibition and hence no transgression; but when reason began its business and weak as it is, got into a scuffle with animality in its whole strength, then there had to arise ills and, what is worse, with more cultivated reason, vices, which were entirely alien to the condition of ignorance and hence innocence. The first step out of this condition, therefore, was on the moral side a *fall*; on the physical side, a multitude of ills [...] The history of *nature* thus begins from good, for that is the *work of God*; the history of *freedom* from evil, for it is the *work of the human being*. (MAM 8: 115)

And in the Concluding Remark, Kant draws on the implications of the last sentence to deploy the third strategy he refers to in "Miscarriage," i.e., the shifting of responsibility for evil away from God toward human beings:

The thinking human being feels a sorrow, one which can even become a moral corruption, of which the thoughtless knows nothing: namely, discontent with the providence that governs the course of the world on the whole, when he estimates the ills that so much oppress humankind, and (as it appears) leaves it with no hope for anything better. But it is of the greatest importance to be *content with providence* ... partly in order to grasp courage even among toils, and partly so that by placing responsibility for it on fate, we might not lose sight of our own responsibility, which perhaps might be the sole cause of all ills, and avoid the remedy against them, which consists in self-improvement. (MAM 8: 120-1).

These claims are not outliers: "Idea for Universal History" (1784) is also a page torn from the *Lectures*. Propositions One to Four present a similar evolutive account (though now coached in a more confident epigenetic tone), while the rest of the essay conceives of competitiveness and antagonism as the schoolmasters of humanity (Propositions Five to Nine). However, nowhere does Kant show his theodicean proclivities more artlessly than in these sentences:

Thanks be to nature, therefore, for the incompatibility, for the spiteful vanity, for the insatiable desire to possess or even to dominate. For without them all the excellent natural predispositions in humanity would eternally slumber undeveloped. The human being wills concord; but nature knows better what is good for the species: it wills discord. He wills to live comfortably and contentedly; but nature wills that out of sloth and inactive contentment he should throw himself into labor and toils, so as, on the contrary, prudently to find out the means to pull himself again out of the latter. The natural incentives to this, the sources of unsociability and thoroughgoing resistance, from which so many ills arise, which, however, impel human beings to new exertions of their powers ... thus betray the ordering of a wise creator, and not the hand of an evil spirit who might have bungled this splendid undertaking or ruined it in an envious manner. (laG 8: 21-22)

In all these cases, Kant engages in what he would later identify as "doctrinal theodicy" (MpVT 8: 264), for he interprets our widespread *experience* of the counterpurposive in the world (sin, pain, and injustice) according to a providential conception of history (i.e., a divine arrangement of events) in which their counterpurposiveness is only apparent – a nuisance necessary to speed up our moral progress. To attain global intelligibility for human history, Kant does not hesitate to make constitutive assertions about what "nature" wills (laG 8:19). Despite some isolated disclaimers (laG 8: 29, MAM 8: 109-10) and gestures toward moral hope, the tone of Kant's writings on history is predominantly apodeictic and does not hide the pretension to

have “*insight into the relationship in which any world as we may ever become acquainted through experience stands with respect to the highest wisdom*” (MpVT 8: 263, italics in the original).²⁰

Now, if we adopt the vantage point of *Miscarriage* to read these earlier works, it is easy to see what the late Kant must have found wrong with them. As Gressis (2018) and Duncan (2012) rightly indicated, doctrinal theodicies lead to the moral evisceration of evil, which is presented as an evanescent inconvenience, a transitory obstacle designed by God to spur our moral progress. For this “ill,” however, we cannot be blamed: immorality results from inescapable “limitations of humanity as a finite nature” (MpVT 8: 259). Soon Kant will discard the *privatio boni* doctrine that undergirds this line of thought and replace it with a *positive* conception of evil, according to which immorality results from an affirmative human choice:

The human being must make or have made himself into whatever he is or should become in a moral sense, good or evil. These two [characters] must be an effect of his free power of choice, for otherwise they could not be imputed to him and, consequently, he could be neither morally good nor evil. (RGV 6: 44)

But this shift regarding the concept of evil reflects a deeper problem: behind these opposing interpretations of immorality lie also incommensurably different conceptions of God. For, in the theodicean accounts of the 1780s, God (often in the guise of providence) appears as a great manipulator, willing to sacrifice individuals for the collective advancement of His plan.²¹ The logic of historical progress is ruthless: the upward trajectory of the species is paved by its members’ fallenness (MAM 8:115). This is so because cultural and moral development require the systematic violation of the individuals’ desires and intentions, i.e., the sacrifice of concord and comfort for the sake of war and toil, the instruments God uses to unleash our creative forces (IaG 8:21-22). In this progressive tale, divine and human interests are inevitably at odds: humans are created in self-contradiction, torn between reason and instinct, sociable and selfish tendencies, forcing God to use extreme measures to advance His vision. A God like that, however, acts in ways incompatible with Kant’s practical philosophy, which is committed to the view that “the human being and in general every rational being exists as an end in itself, *not merely as a means* to be used by this or that will at its discretion” (GMS 4:428).²² To avoid instrumentalizing humanity, a *moral* Creator would have to will in such a way that human beings be able to contain within themselves the plan for His creation. However, for the longest period, the fledgling character of human rationality precludes this possibility. Only creatures who have been released from the despotism of nature and are in full control of their rational capacities would be able to contain within themselves God’s ends without violating their freedom (GMS 4:429).

Thus, in ways the late Kant would have found scandalous, the God of doctrinal theodicies resembles the God in the Biblical story of Job. This is a God who submits the most faithful of his servants to utter despondency and humiliation, who acts behind Job’s back and engages in a plan that deserves to be called “gaslighting” –all to prove to Satan that Job will remain faithful.²³ There are no traces of such cruelty in *Miscarriage*, where the story is framed *exclusively* for our moral edification, no matter how much hermeneutic violence this framing might introduce into the Biblical text. Once again, *Religion* contains the key to understanding what Kant is doing in this essay:

What is being said here must not be regarded as though intended for Scriptural exegesis, which lies outside the boundaries of the competence of mere reason. We can explain how we put a historical account to our moral use without thereby deciding whether this is also the meaning of the writer or only our interpretation, if this meaning is true in itself, apart from all historical proof, and also the only meaning according to which we can derive something edifying from a text which would otherwise be only a barren addition to our historical cognition. We should not quarrel over an issue unnecessarily, and over its historical standing, when, however we understand it, the issue does not contribute anything to our becoming a better human being ... Historical cognition that has no intrinsic relation, valid for everyone,

²⁰ One of the anonymous reviewers of this paper raised objections to my interpretation of Kant’s historical writings. Drawing from some of the same passages I mentioned above (particularly, IaG 8:29), the reviewer emphasized the pragmatic, reflective character of Kant’s judgments about providence’s plan for history. Such passages, along with the openly conjectural approach Kant adopts in *Conjectural Beginning*, prove, in the reviewer’s mind, that my constitutive reading of these texts is off-track. These are not meant as theoretical judgments, the reviewer concludes, but as ideas that can be useful for practical purposes. This binary contrast, however, misses the point. My claim is not that Kant was a thoroughgoing rationalist during this period. It is, rather, that both rationalist and critical tendencies coexisted within him, and that it is his rationalist part that led him to embrace, against his better philosophical judgment, the same doctrinal views he will eventually come to reject. Kant’s texts do not present us with an either/or; they present us, rather, with a troublesome vacillation, a back and forth that indicates the stubborn presence of his rationalism.

²¹ For an elaboration of this view, see my “The Birth of God, and the Problem of History,” *Kant and the Question of Theology*. Chris Firestone, Nathan Jacobs, and James Joiner (eds.). New York: Cambridge University Press, 2017, 35-56.

²² See Barbara Herman, “A habitat for humanity,” in *Idea for Universal History: A Critical Guide*, Amélie Rorty and James Schmidt (eds.). Cambridge University Press, 2009, pp. 150-170, particularly p. 151.

²³ As I read the biblical story, the Book of Job contains perhaps the first written record of “gaslighting.” For, without Job’s knowing, God and Satan, along with Job’s friends who gather with the alleged intention of “consoling” him, conspire to put Job’s sanity into question. This set of circumstances fits to the letter the description of gaslighting Kate Abramson provides: “Very roughly, the phenomenon that’s come to be picked out with that term is a form of emotional manipulation in which the gaslighter tries (consciously or not) to induce in someone the sense that her reactions, perceptions, memories and/or beliefs are not just mistaken, but utterly without grounds—paradigmatically, so unfounded as to qualify as crazy” (p. 2). See Kate Abramson, “Turning Up the Lights on Gaslighting,” *Philosophical Perspectives*, 28 (1), 2014: 1-30.

to this [moral improvement], belongs among the adiaphora, which each may treat as one finds edifying. (RGV 6: 43-4n.)

Seen through this interpretative lens, much of Scripture becomes “adiaphora” and can hence be treated without regard for its literal or historical truth. It is from this treatment that Job emerges, along with Jesus, as a founding father of Kant’s moral religion –not a patriarch of the Old Testament, as the Biblical story has him, but a Christ-like figure who would see the Scriptural version of his life with the same apprehension that Kant must have felt towards his earlier theodicean views. Something similar should be said with respect to the Kantian God, who appears in *Religion* as a purely rational and moral deity after His conception has been rid of the impurities Scripture accrues on it. All these figures have been scrubbed clean, submitted to the stern discipline of critique –Kant himself included.

-V-

If our interpretation is correct and doctrinal theodicies not only undermine human responsibility, but also corrupt the divinity of God in the process, what could have led Kant to entertain them? How could he have so blatantly disregarded the constraints the first *Critique* imposes on any kind of theological speculation?

There is no need to beat around the bush for an answer. Kant gives us a clear indication in *Miscarriage* of where he went wrong in those earlier years. His insight is that, to abandon this kind of project “*once and for all*,” it is not enough to refute theodicean claims “*ab instantia*” (MpVT 8: 263). What we need to do, instead, is to find the fundamental error in this line of thinking, i.e., the principle from which doctrinal theodicies derive, so that “all further attempts by a putative human wisdom to gain insight into the ways of the divine wisdom are fully dismissed” (*ibid.*). The “juridical process” in the first part of the essay, Kant believes, makes this error clear: “our reason is absolutely incapable of insight into the *relationship in which any world as we may ever become acquainted with through experience stands with respect to the highest wisdom*” (*ibid.*). Such a conclusion, however, will sound like a platitude to any reader of the first *Critique*, a text whose declared purpose was to establish the limits of human knowledge. What is remarkable, then, is not what this platitude actually says (a critical commonplace) but what it *shows*, for it is those very limits which Kant must have ignored when he espoused doctrinal theodicies a few years before.

Indeed, Kant’s ignorance was so complete that he now identifies remembering our noumenal ignorance with “negative wisdom.” This identification is appropriate. For, if we consider the general etiology of error, we can see that Kant’s forgetfulness is not a simple lapse of memory –it is, rather, a *deed*, an accomplishment of sorts. This is so because, as Kant tells us in the Jaesche *Logic*, to err “a falsehood” must be “taken for [a] truth” (Log 9:53). This “taking as” is an act of judgment for which we are responsible: “An erroneous judgment– for there is error as well as truth only in judgments –is thus one that confuses the illusion of truth with truth itself” (*ibid.*). This confusion, Kant believes, “is effected only through the **unnoticed influence** of sensibility on understanding, through which it happens that the subjective grounds of the judgment join with the objective ones ... and make the latter deviate from their destination” (KrV B551/A294, my emphasis). For this “deviation” to be possible, it follows, the distorting “influence” of sensibility must take place below the threshold of consciousness, for a falsehood recognized to be such could not possibly persuade us to adopt it. Error, therefore, emerges only when a falsehood disguises itself to us as truth and produces an illusion imperceptible to the one who holds it.²⁴

Our collusion with error is crucial, for “[t]he will does not have any influence immediately on holding-to-be-true” (Log 9:73); otherwise, “we would constantly form for ourselves chimeras of a happy condition, and always hold them to be true, too” (Log 9:74). But this lack of influence is not absolute: even if wish and reality cannot be interchanged as one pleases, our will has the power to blur their boundaries. Although it does not do so *directly* (since it cannot force our assent to a falsity we recognize as such), it has plenty of leeway to resort to *indirect* tactics (Log 9:74).

Throughout the corpus, Kant calls the indirect influence of the will on beliefs “rationalizing” (“*Vernunfteln*”), a notion meant to designate the subterfuges our reason uses to muddle the difference between truth and falsehood, objective and subjective grounds –all in ways that are unnoticed by the agent and nonetheless culpable, for the misuse of one’s faculties could not take place without one’s active involvement and cooperation.²⁵

Now, although critical discipline plays an essential role in developing relevant countermeasures, its efficacy is often impaired by the fact that the subjective grounds at the basis of error do most of their work *incognito*, undetected by introspection. To deceive us, what we want to believe must be able to present itself as a truth, while at the same time erasing any trace of such deception, for otherwise our epistemic defenses would not be breached. As Kant sees it, our complicity in their infringement lies at the heart of error and places critical vigilance and our rationalizing tendencies in a constant tug of war (KrV A295/B351).

According to this model, the stronger the pull of the subjective grounds, the more stringent the vigilance must be to counteract it. But vigilance is a *conscious* activity, and hence always arrives *after* rationalizing has done its work. Kant considers transcendental illusions as a paradigmatic example of such a protracted response, for they “[do] not cease even though [they are] uncovered and [their] nullity is clearly seen by transcendental criticism” (KrV A297/B353). Unlike common errors, which can be cleared away once their falsity

²⁴ I argue at length for this view in “Emotivism: A Kantian Response,” *Society of German Idealism Review (SGIR)*, 6.1 (2023), 60–82.

²⁵ It is worth noting that Kant uses a variant of this word (“*Vernünfteln*”) in *Miscarriage* to characterize the spurious use of speculative reason that gives rise to doctrinal theodicies (MpVT 8:264).

has been exposed (Log 9:56), the persistence of transcendental illusions is due to the centrality of the subjective interests they come to satisfy. In their case, it is not an idiosyncratic preference of the subject, but her desire to *know* the unconditioned –a desire Kant considers “inextricable” for a finite reason like ours:

The cause of this is that in our reason (considered subjectively as a human faculty of cognition) there lie fundamental rules and maxims for its use, which look entirely like objective principles and through them comes about that the subjective necessity of a certain connection of our concepts on behalf of the understanding is taken for an objective necessity, the determination of things in themselves. (KrV A297/B353)

Although “this is an **illusion** that cannot be avoided at all,” for it is “one that irremediably attaches to human reason ... continually propelling it into momentary aberrations that always need to be removed” (KrV A297-8/B354), its deceptive power can be neutralized if we learn to coexist with the temptation and respect our cognitive limits. This is where the discipline of reason comes in play: the self-knowledge it entails provides the training necessary to increase our capacity to detect and resist self-deception.

My claim in this paper has been that this is the training Job’s reason must have undergone but that the earlier Kant failed to undertake, for a part of him was still committed to the rationalist assumption that metaphysics was first philosophy. My contention now is that he overlooked the problem for so long because he fell prey of a transcendental illusion of sorts. For doctrinal theodicies, like transcendental illusions, are the product of speculative self-conceit and their falsehood remains undetected to the one who holds them. Critical vigilance is doomed to stay a step behind in this process. As a conscious activity battling an enemy that hides behind reason, vigilance can at most do *damage control*: it can protect us against the worst consequences of the deception but arrives only after the deception has taken place.

In the case of doctrinal theodicies our impotence to resist the illusion is compounded. For the temptation that gives rise to them combines the promise of ultimate intelligibility (as all transcendental illusions do) with the need to contrive a positive conception of oneself (as the friends of Job illustrate). It is the combined effect of these powerful subjective forces, the “illusion of truth” mingled with the “illusion of goodness,” which sets in motion the rationalizing tendencies that give rise to peculiar kind of error at the heart of philosophical theodicies.

As Kant portrays them, Job’s friends interpret their well-being as a sign of divine favor, and thereby confuse their happiness with their virtue, forming an image of God’s justice “according to [their] station” (MpVT 8: 265). With this image of the divine, they feel themselves justified in their own probity, for their sense of moral worth is obtained by inverting the order of priority between God’s attributes: they make “benevolence,” i.e., the secondary attribute that pertains to God as a “ruler” (MpVT 8:257), “the supreme condition of world creation to which the holiness of legislation is subordinated” (MpVT 8:257 n.). By neglecting the primacy of God as “lawgiver,” Kant argues, they do “violence to religion, which has this very concept for foundation” (*ibid.*). But their violence is not accidental: it is necessary to satisfy an essential subjective need, for “[i]ndeed the human being wishes to be happy first” (*ibid.*).

Unlike Job, his friends refuse to accept “that the worthiness to be happy, i.e., the conformity of the employment of [one’s] freedom with the holy law, must in God’s decision be the condition of his beneficence, and must therefore, necessarily precede it” (MpVT 8: 257 n.). Complacent in their good fortune, they fancy a God who reflects onto them what they want to believe, and hence “declare themselves for the system which explains all ills in the world from God’s *justice*, as so many punishments for crimes committed” (MpVT 8:265). This view of the divine is an expression of their heteronomous self-conception, according to which rewards and punishments are the primary motivators. Based on this conception, they construe a picture of the divine that validates who they think they are, creating a “religion of supplication” (MpVT 8:267) in which God acts as a benevolent ruler that rewards their good behavior.

It is not difficult to see at the root of this spurious religious conception evidence of the same “natural dialectic” Kant mentions at the end of *Groundwork I*. For, confronted with the unrelenting demands of morality, the friends of Job (like most of us) set in motion:

a propensity [*Hang*] to rationalize [*vernunfteln*] against those strict laws of duty and cast doubt on their validity, or at least on their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis, and to destroy all their dignity ... (GMS 4: 405)

My point is that, along with the “illusion of truth” which gives rise to transcendental illusions, the “illusion of goodness” generates, in the blurry eyes of their practitioners, the mendacious undertaking Kant calls “doctrinal theodicy” in *Miscarriage*. What is curious, although not surprising once we understand the crooked psychology of error, is that it took Kant a whole decade to recognize the presence of these devious stratagems within himself. Once this recognition took place, he was a step closer to give this duplicity a new name: he came to all it “the radical evil in human nature” (RGV 6:19).

-VI-

I realize that the interpretation I am proposing may strike some readers as at best conjectural, and at worst implausible. I am guilty as charged on the first count but not on the second, for I provided ample textual evidence in support of this view and resorted to conjecture only to elucidate a scandal Kant refused to acknowledge, namely, the scandal that for almost a decade he entertained theodicean projects that contradicted his

best philosophical judgment. This perplexing behavior, I have argued, is akin to the type of motivated irrationality Kant identified with “transcendental illusions,” a form of error we cannot avoid and discover through its consequences (in this case, the curtailing of our moral responsibility and the desecrating of the divinity of God). When it comes to doctrinal theodicies, the error is due to two inextricable needs of human reason: the speculative need to know the unconditioned (God), which it shares with traditional transcendental illusions, and the practical need to think well of ourselves, which characterizes this dubious religious project. I maintained that these needs give rise to two complementary types of illusion, the “illusion of truth” and the “illusion of goodness.” Combined, they lead us to imagine a world where the laws of nature and the laws of freedom are perfectly harmonious, the creation of a benevolent God that arranges the universe with our happiness in mind. In such a world, the counterpurposive can be wished away, for evil and suffering are seen as temporary inconveniences, tools that God uses to speed up our moral progress. In *Miscarriage*, Kant associated this questionable worldview with the perspective of Job’s friends, whose consolations he sees as symptomatic of a deep form of mendacity, a cunning way to present themselves as defenders God’s cause. Throughout the 1780s, I have shown, Kant was not immune to the appeal of this deception.

Read in this light, the *Theodicy* essay can be interpreted as an overdue act of self-criticism, a late attempt of the critical Kant to release himself from the grip of his own rationalism. The effects of this release will soon be visible in *Religion within the Boundaries of Mere Reason*, where self-deception is no longer associated with archetypal individuals, but becomes “the foul stain of our species” (RGV 6:38). Universalizing mendacity, I suspect, must have been for Kant the only way he was able to accept his own.

Bibliography

- Abramson, Kate. “Turning Up the Lights on Gaslighting.” *Philosophical Perspectives* 28, no. 1 (December 2014): 1–30. <https://doi.org/10.1111/phpe.12046>.
- Brachtendorf, Johannes. “Kants Theodizee-Aufsatz – Die Bedingungen Des Gelingens Philosophischer Theodizee.” *Kant Studien* 93, no. 1 (January 6, 2002): 90–102. <https://doi.org/10.1515/kant.93.1.57>.
- Cavallar, Georg. “Kants Weg von Der Theodizee Zur Anthropodizee Und Retour: Verspätete Kritik an Odo Maquard.” *Kant Studien* 84 (1993): 90–102.
- Duncan, Sam. “Moral Evil, Freedom, and the Goodness of God: Why Kant Abandoned Theodicy.” *British Journal for the History of Philosophy* 20, no. 5 (September 2012): 973–91. <https://doi.org/10.1080/09608788.2012.718860>.
- Gressis, Robert. “Kant’s Theodicy and Its Role in the Development of Radical Evil.” *Archiv für Geschichte der Philosophie* 100, no. 1 (March 7, 2018): 46–75. <https://doi.org/10.1515/agph-2018-0003>.
- Herman, Barbara, and James Schmidt. “A Habitat for Humanity.” Essay. In *Idea for Universal History: A Critical Guide*, edited by Amélie Rorty, 150–71. Cambridge University Press, 2009.
- Huxford, George. *Kant and Theodicy: A Search for an Answer to the Problem of Evil*. Lanham, Maryland: Lexington Books, 2020.
- Larrimore, Charles. “Autonomy and the Invention of Theodicy.” Essay. In *New Essays on the History of Autonomy: A Collection Honoring J.B. Schneewind*, edited by Natalie Brender and Larry Krasnoff. Cambridge University Press, 2004.
- Muchnik, Pablo. “Emotivism: A Kantian Response.” *Society of German Idealism Review* 6, no. 1 (2023): 60–82.
- Muchnik, Pablo. “The Birth of God, and the Problem of History.” Essay. In *Kant and the Question of Theology*, edited by Chris Firestone, Nathan Jacobs, and James Joiner, 35–56. Cambridge University Press, 2017.
- Pasternack, Lawrence. “Kant’s Moral Argument and the Problem of Evil: Authentic Theodicy and the Sincerity of Faith.” *Kant on Proofs for God’s Existence*, December 18, 2022, 139–60. <https://doi.org/10.1515/9783110688962-009>.
- Schulte, Christoph. “Zweckwidriges in Der Erfahrung. Zur Genese Des Mißlingens Aller Philosophischen Versuche in Der Theodizee Bei Kant.” *Kant Studien* 82, no. 4 (1991): 371–96. <https://doi.org/10.1515/kant.1991.82.4.371>.
- Seeskin, Kenneth. “Job and the Problem of Evil.” *Philosophy and Literature* 11, no. 2 (October 1987): 226–41. <https://doi.org/10.1353/phl.1987.0056>.