

Kant on Negative Wisdom: Faith and the Counterpurposive in “On the Failure of All Attempted Philosophical Theodicies”

Thomas Scott Ferguson
Independent Scholar 

<https://dx.doi.org/10.5209/kant.105213>

Received: 16-10-2025 • Accepted: 16-12-2025

ENG Abstract: This article rereads Kant’s difficult 1791 essay on theodicy as constructing a challenge to Kant’s own critical system, especially the end of third *Critique*, in regard to the compatibility of nature and morality and (by extension) of speculative and practical reason. He uses this challenge to redevelop his concept of wisdom, playing down the systematic/teleological role of previous texts and presenting a novel, *apophatic* mode of wisdom based in sincerity. I argue that authentic theodicy (illustrated in Kant’s reading of Job) “handles” sincerely held doubts by revealing to the faithful moral agent the “inscrutability” of a creation which nevertheless remains wise, and at the same moment disclosing those agents’ own status. Doubts against divine wisdom are not so much answered as contextualized within “good life conduct.” The consequence is a retreat in the farthest ambitions of Kant’s system (the unification of reason), and increased emphasis on self-cognition.

Keywords: Immanuel Kant; Theodicy; Philosophy of Religion; Job

Summary: Introduction: A Counterfactual Kantian Theodicy. 1. Wisdom, Teleology, and Systematicity. 2. Counterpurposiveness and the Fall of Righteous Spinoza Theodicy. 3. Why Negative Wisdom? 4. Negative Wisdom I: Initial Presentation and Objections. 5. Negative Wisdom II: Authentic Theodicy and Job. 6. The Phenomenology of Counterpurposiveness. 7. Sincerity and Critique. 8. Negative Wisdom and Reflective Faith. 9. Conclusions

How to cite: Ferguson, Th. S. Kant on Negative Wisdom: Faith and the Counterpurposive in “On the Failure of All Attempted Philosophical Theodicies”. *Con-textos Kantianos. International Journal of Philosophy*, 23, 61-77. <https://dx.doi.org/10.5209/kant.105213>

Kant’s one dedicated essay on theodicy is a confounding text, even for Kant veterans. In this paper I attempt to clarify this essay by reading it as a redevelopment of Kant’s systematically crucial concept of “wisdom” and a reinterpretation of his own concept of faith. Before I introduce my argument, however, I want to articulate why this essay is so confounding.¹

Introduction: A Counterfactual Kantian Theodicy

In the decade following the *Critique of Pure Reason* Kant’s works consistently argue that, to the extent that reason cannot answer fundamental questions concerning the objects of metaphysics through either argumentation or intuitive speculative insight, it can nevertheless reinterpret those objects as practical presuppositions necessary to render possible the achievement of reason’s final end. This end is a “moral world” (KrV, A808/B836) or “realm of grace” (KrV, A812/B840) in the first *Critique*, later the “highest good.”² Where speculation falls

¹ I am grateful to countless interlocutors for their help in the course of my drafting this paper, and I particularly thank Andrew Chignell, Pablo Munchik, Lawrence Pasternack, and the anonymous reviewers of *Con-Textos Kantianos* for their comments and suggestions. In citing Kant I follow the standard practice of citing abbreviations and paginations from the Akademie Ausgabe, with the exception of the first *Critique* (where citations follow the A and B editions). Translations are taken from the Cambridge Edition and amended as necessary.

² Nothing in this part of the scholarship should be taken for granted, and the elision of the “moral world” of the first *Critique* with the “highest good” of the second is not uncontroversial—Reath, for example, distinguishes the moral world as a “secular conception” (Reath 1988, 615) against what he sees as a more theological, otherworldly, and problematic conception of the highest good. I do not think this is right: Kant puts considerable emphasis on the ultimate compatibility between nature and morality, and for me that is enough to suppose that the highest good in the world (as the final end of reason) should not be taken as referring to a world any other than that nature. The details of Kant’s conception change considerably between 1781 and 1788 and I sidestep those details for the purposes of this paper; but if I rely more heavily upon the earlier language of a “moral world,” it is because I wish to empha-

short, the argument goes, rational faith grounded in practical reason still offers metaphysics a “source of positive cognition” (KrV, A795/B822) for God, the immortal soul, and the world, with a legitimate though merely practical sense. Kant’s principal way to show that and how reason still provides a response to the questions that endlessly burden it, and at the same time dismiss the pretensions of dogmatists and theosophers as well as skeptics and naturalists, is his distinction between the certitude of *knowledge* and the certitude of *faith*.³

The reader of Kant might therefore expect that a distinctively critical approach to *theodicy*, one of the classic topics of modern philosophy after Leibniz, would deploy a similar line of argument.⁴ And indeed, one can identify just such a nascent critical theodicy within a well-known passage at the end of the 1790 *Critique of the Power of Judgment*: the last paragraph of §87, while ultimately a variation upon Kant’s moral proof for the existence of God, also directly addresses the problem of evil and purposelessness in the world. Although taking up only one paragraph, it indeed deploys the strategy above—namely transposing a metaphysical question from a theoretical key to one of practice and faith—to deal with that problem, and so represents what a capable reader of Kant might have expected a “critical treatment” of theodicy to look like.⁵ “[A] righteous man (like Spinoza),” Kant says, “who takes himself to be firmly convinced that there is no God and... there is no future life” can still very well work to “unselfishly establish the good to which the holy law directs all his powers” (KU, AA 5:452), but he always finds himself confronted with a world full of ills and injustices. More than that, he stands before the “abyss of the purposeless [*zwecklos*] chaos of matter,” with the purposelessness being the real problem. Without any apparent moral purpose to nature, Kant says, the efforts of this “Spinoza”⁶ must remain “limited [*begränzt*].” If this character can accept that a moral author of the world is “at least not self-contradictory” (KU, AA 5:453), however, then he has good reason to become a faithful theist “from a practical point of view.”⁷ If the right kind of God is at least possible, it renders his moral ends equally possible and removes his limitation.⁸

size that Kant’s highest good is a *this-worldly* one—albeit one that still calls for a God. See Beiser 2006 and, more recently and comprehensively, Höwing 2016.

³ Faith (*Glaube*), for Kant, is a mode of holding-for-true in some ways overlapping with “believing” in the everyday sense, but also a distinctively practical kind of comportment. “Once an end is proposed,” writes Kant, “then the conditions for attaining it are hypothetically necessary” (KrV, A823/B851)—which is to say that once I *will* something, then I must hold as true the means and other conditions which make the realization of that goal possible. If a doctor gives medicine to a sick patient in order to cure them, their relation to the medicine is not that of theoretical inspection or judgment; rather they *rely upon it* in willing their end, and Kantian faith is just that practical reliance. Such reliance indeed involves a number of contingent theoretical stances—concerning what the patient is sick with, the effectivity of the medicine, etc.—but there is no faith proper until I use what I know and act, until I rely on something in the “actual use of the means to certain actions” (KrV, A824/B852). Kant then extends this concept to metaphysics. For when willing pursues of the final ends of reason, Kant argues, so it must rely upon—have faith in—the classical objects of metaphysics, however inaccessible those objects may be to any theory. It is this deeper conceptual structure, and not a traditional fideism, that lies beneath Kant’s famous statement that he “had to deny **knowledge** in order to make room for **faith**” (KrV, Bxxx). Wood 1970, especially 16–25, is still an excellent text on this topic; also see Beiser 2006, Chignell 2007, and Stevenson 2003 as good points of reference.

⁴ It is notable that the theodicies presented by Kant in lectures earlier in the 1780s (V-Phil-Th/Pöhlitz, AA 28:1076ff) also do not make considerable use of this strategy, so that they may be just as contrary to the reader’s expectations for a “Kantian theodicy” as the “Failure” essay. See Huxford 2020 49–54 for a strong case for the agreement of these lectures with Kant’s stated views at the time, and 79–90 for a detailed discussion of Kant’s arguments there. Also see Schulte 1991 on the clash between the “Failure” and these earlier texts, as well as Pablo Muchnik’s contribution to this volume; Duncan 2012 offers a case that this clash may stem from Kant’s encounter with Carl Christian Erhard Schmid’s *Moralphilosophie* (published the same year as the third *Critique*), although I share similar skepticism to Kravitz 2020 regarding how decisive this was. On this topic in particular it is important to consider that there is no one completely consistent critical Kant, as treatments of “evil” from the *Lectures* to the third *Critique* to the “Failure” and beyond show enormous change in Kant’s thought on the subject in the space of less than a decade.

⁵ It was a draft of Lawrence Pasternack’s last paper, now published in this volume, which first suggested to me that this passage could be read as a distinctively Kantian theodicy, and I build the “RST” below on the basis of his insight.

⁶ For this paper I keep to Kant’s word when discussing the protagonist of this passage and related arguments, while also using the full title “righteous Spinoza” (or setting “Spinoza” in quotations) to make it clear that it is a *Kantian character* who is intended and not the man. Kant’s writings contain two characters both named “Spinoza,” these characters dating back at least to the first success of Bayle’s 1697 *Dictionnaire Historique et Critique*. The third *Critique* features the longest engagement with Spinoza the thinker that Kant published in his lifetime, especially at KU, AA 5:391–95 and in passing at KU, AA 5:417–421; I know of no definitive evidence that Kant ever studied Baruch Spinoza’s writings directly rather than receiving him secondhand through Bayle, Jacobi, and others, but Kant’s engagement with “Spinoza the thinker” nevertheless becomes increasingly important from 1790 onwards (see Allison 1980). The second character is Spinoza as *atheist* (however far Baruch Spinoza himself would not qualify as atheistic by present standards) who nevertheless lives virtuously, a scandalous thought for the time, and it is this “righteous Spinoza” who is deployed as a case study in the passage above. However, although the particularities of Baruch Spinoza’s actual life and thought are unnecessary for Kant’s point—Spinoza is not even the only thinker Kant responds to in this way (cf RezSchulz, AA 8:13–4 as a further example)—this name is not chosen without reason. Spinoza the thinker first comes up at AA 5:391–95 as a representative of fatalism, i.e. in his rejection of final causality, while this passage featuring Spinoza the righteous atheist is precisely about the practical need for such final causality (and ultimately for the moral faith necessary for it). To all appearances, then, the issue of the final causality (teleology) of nature is why Kant picked a Spinoza for this role over just any righteous nonbeliever, so that here at least these characters are quite related.

⁷ Although only discussed in later pages, I take it as crucial for this argument that Kant thinks that only “critical insight into the limits of speculative reason” can render the righteous Spinoza’s *Zweifelglaube* into practical affirmation (KU 5:472–3). The *Collins Ethics* is much more explicit about the nature of this “Spinoza’s” problem, which Kant sees only as excessive emphasis on the speculative: “Atheism can reside in mere speculation, while in practice such a person can be a theist or venerator of God, whose error extends to theology, but not to religion. Those persons who have fallen into atheism through speculation should not be so readily condemned for wickedness as the custom is; only their understanding was corrupted, not their will; Spinoza, for example, did what a man of religion should do. His heart was good, and could easily have been brought to rights; he merely had too much trust in speculative argument” (V-Mo/Collins, AA 27:312). Kant seems to think that the character he describes is already positioned for moral faith, and only requires the removal of speculative errors.

⁸ Most commentators, including Pasternack, endorse psychologically-motivated readings of this “righteous Spinoza” passage. Even readers like Chignell, who take the moral proof of the second *Critique* (KPV, AA 5:124–5) as centrally concerning “rational in-

Assuming familiarity with Kant's treatment of rational theology and faith, then, all of the above is abundant material for a more structured theodical argument. It could go like so:

1. With the critique of speculative theology having ruled out all claims about God made beyond the bounds of human knowledge (that God does or does not exist, or that God is or is not moral, etc.), still the *real possibility* of the existence of a moral author of the world remains in play.
2. A maximally holy, good, and just creator of the world would not have created any absolute evils or rendered nature absolutely purposeless, so that the assumption of such a God also implies assuming that any apparent evil or purposelessness in the world is at least not totally incompatible with the final end of reason.
3. The morally committed agent risks rational and/or psychological instability if they must give up their final end as impossible.
4. Absolute evil or purposelessness in the world would render the final end of morality impossible.
5. Even if speculative certainty is impossible, at a *practical* level the agent still has need for faith in a moral author of the world (on pain of instability and its consequences) at least from a practical standpoint.

Let us call this kind of argument a Righteous Spinoza Theodicy, or RST. Its chief feature is that practical faith in the right kind of God, in exhibiting the world as "purposive," is deployed to render the evils of the world endurable and extend the limited moral projects of the agent without appeal to any speculative arguments concerning God, evil, or the future. Although the "problem of evil" here may be the tension between an amoral world and final moral ends rather than between evil and divinity, the ultimate resolution—obtaining the correct concept of God (as moral author) and, by extension, of the world (as purposive)—is thoroughly theodical. And if Kant chooses a morally committed atheist as his subject, still in principle the argument would apply to any agent who does not assume the right kind of theological situation (e.g., one who believes in a God but denies their goodness).

Now because such a RST is given a prominent if underdeveloped place in the third *Critique*, we could have expected that Kant's "On the Failure of All Attempted Philosophical Theodicies"—published the following year—would follow its form. But much of what makes this essay confounding, to return to my opening, is that it does not do this. The "Failure" does assume such arguments as a background in important respects—making it key for interpretation—but it is very far indeed from being an RST itself. It instead reads as an assault on the very project of theodicy, where faith in God does not relieve us from the abyss but instead, in the climax of the Book of Job, throws the faithful man straight at it. If this text is confounding, it is an indication that the problems Kant addresses, and Kant's reasons for writing it, demand further clarification.

This paper attempts such a clarification in several steps, reading the essay as a novel reconception of Kant's concept of *wisdom* which goes in tandem with an equally novel reinterpretation of the dynamics of moral faith. I first argue that the 1791 essay presents "the problem of evil," under the term *counterpurposiveness*, in a way specifically constructed by Kant to challenge the coherence of his own critical system (and the RST with it); Kant's own sense of the term "wisdom" is crucial to understanding the theodical plaintiff's challenge and the direction of the rest of the essay. Kant's remarkably sympathetic response to the doubts of these plaintiffs then requires his revision of the farthest ambitions of that system, particularly through a distinction in the concept of wisdom between divine wisdom and what Kant calls *negative* wisdom. Kant rescues the compatibility of nature and morality by rendering that compatibility—and thus the whole of creation—"inscrutable," with such inscrutability being a privilege and not just a shortcoming of our own knowledge. Yet the more important aspect of this "authentic" theodicy, as Kant presents it through his reading of Job, is that it is supposed to allow the faithful moral agent to make sense of themselves within that creation. Doubts show themselves, on reflection, to be not theological problems to be corrected but symptoms of sincere faithfulness, so that theodical issues are to be resolved through a moral life rather than argument or coercion. At the same time, however, the newly central role of sincerity in negative wisdom also returns us to basic methodological themes of critical philosophy, namely self-cognition.

stability" (to use John Hare's phrase), have at least two good reasons to take this 1790 passage as distinct from that of the KPV: one is that Kant's language *prima facie* sounds like that of psychological need, and another is that such a variation on the argument is potentially more compelling (Chignell 2023, 50-54). I briefly mention two misgivings regarding such readings. The first lies in the very existence of the theodicy essay of the following year (the central text of this essay), which—so I argue—goes out of its way to problematize faith as a source of comfort. Secondly, on a psychological reading it is unclear why faith *in particular* must play this role. Consider an argument from David Sussman: beginning from the assumption that "Kant ... suggests that faith serves only to psychologically buttress failing moral resolve," he can neatly conclude that "insofar as the issue here is purely one of motivational effectiveness, then any of the wide range of fantasies might have morally salutary effects, given the right background psychology" (Sussman 2010, 144). Sussman does not consider this "wide range" (he specifically proposes love) a problem; he is himself one of the psychologically driven readers. But it is a broad issue with this interpretation that such a range of sufficient motivators is just not to be found in Kant's texts. The psychological reader can very well bite this bullet and suggest that Kant was short-changing himself—that love and other such "fantasies" should have been options as well—but if they assume that Kant was not oblivious to these options, and they wish to stick to Kant's word, then they need an account of why moral faith in particular plays the role that it does. I suspect that must push them back towards a "rational stability" framework. Although I cannot develop an argument here, I myself lean towards reading this "Spinoza's" limitations in terms of a rational inability to understand themselves and their own projects. Yet one does not thereby drop the psychological side of things: to the contrary, this can provide the psychology a deeper context. When Andrea Marlin Esser suggests just such a heterodox reading of this passage, that "[The atheist's] endeavour is circumscribed because, given the particular way in which he sees the world and himself, he is incapable of appropriately articulating 'his own inner purposive determination by the moral law'" (Esser 2016, 247), it makes good sense that psychological peril can *follow* from such a failure of self-understanding.

1. Wisdom, Teleology, and Systematicity

The very first line of the “Failure” essay already reframes the question of theodicy, if the reader follows Kant’s terms. It reads:

By “theodicy” we understand the defense of the highest wisdom of the creator against the charge which reason brings against it for whatever is counterpurposive in the world. (MpVT, AA 8:255)

Albeit easy to pass over, this first line is coded with crucial jargon—especially “Wisdom” [*Weisheit*] and “counterpurposive” [*zweckwidrig*]—which spells out the systematic stakes of the whole essay.

On its face there seems to be no reason for Kant to write the ornate phrase “highest wisdom of the creator” instead of just God—but he insists on the substitution, even rejecting and lightly ridiculing the traditional language of “defending God’s cause” in the very next sentence. For most readers this would have been a distinction without a difference—who else would the “highest wisdom” indicate, after all?—and Kant certainly doesn’t intend any entity *other* than God. But wisdom also has a very specific sense for Kant, one with a long history in the critical corpus and an architectonic importance well out of proportion to the amount of space he devotes to actually writing about it.

As a term of art, wisdom directly relates to the “cosmopolitan concept” of philosophy discussed in the “Architectonic” of the 1781 *Critique*. The relevant passage reads:

[P]hilosophy in a genuine sense ... *relates everything to wisdom* ... Mathematics, natural science, even the empirical knowledge of humankind, have a high value as means, for the most part to contingent but yet ultimately to necessary and essential ends of humanity, but only through the mediation of a rational cognition from mere concepts, which, call it what one will, is really nothing but metaphysics. (KrV, A850/B878, my emphasis)

Similar passages are to be found throughout the corpus, albeit rarely with extended discussion.⁹ If philosophy is “the science of the relation of all cognition to the essential ends of human reason” (KrV, A839/B867), then wisdom represents *achievement* of this relation, the teleological alignment of everything towards the final end of practical reason.¹⁰ Wisdom’s role manifests early in the critical period, consistently accompanied by the concept of *Zusammenstimmung*—literally “tuning together” everything towards reason’s ultimate ends. Fragment 4849, dated by Adickes to 1776–9, presents a sketch of the critical system where wisdom—“the agreement [*Zusammenstimmung*] to the sum of all ends”—plays a role as the core of the “complete *principium* of practical unity”¹¹ which metaphysics must demonstrate. The completion of the business of our reason, Kant says a few lines down, is “an agreement of theoretical with practical unity, where the practical use [of reason] is led beyond the boundaries of the pragmatic, hence also beyond the present world ... in relation to the conditions of a complete unity” (AA 18:6).¹² But this unity between the two uses of reason, Kant says, is always a question of “tuning together” all cognition towards practical reason’s final end. Wisdom *just* is this idea of a completed unity of reason, the teleological realization of both knowledge and will with the two working in synergy. This gives wisdom a deep systematic importance: since the practical and theoretical applications of reason have distinct interests¹³ which can set them at loggerheads, wisdom names the critical goal of finally putting those applications into harmony.

It is crucial for understanding the “Failure” essay that wisdom and its teleological unification of practical and theoretical reason call for *the right kind of world*, a world which reason can interpret as amenable to morality. Kant is clear that practical reason’s hope for actually achieving its final end structurally relies upon theoretical reason’s capacity to find means within the world adequate to such an achievement. What ultimately “unifies practical with speculative reason,” Kant says in the first *Critique*, is the idea of the moral world *leading into* the “purposive unity of all things” under “universal laws of nature,” which are both united by the presupposition of an intelligent creator (KrV, A815/B843).¹⁴ Reason therefore has an interest in viewing the order of nature as having arisen from a “moral order”:

⁹ *Weisheit* is most commonly discussed in its relation to philosophy as the “doctrine” of wisdom; beyond the passages quoted here, see (as good examples) KrV, A850/B878 and A838–9/B866–7, KpV, AA 5:108, MS, AA 6:375 and 6:441, PKR, AA 8:441, Log, AA 9:23–26, V-Lo/Dohna, AA 24:698, and V-Phil-Th/Pöhlitz, AA 28:1057. The lack of extended discussion in the corpus, i.e. its scattered and decentralized status, is also reflected by a relative dearth of scholarship on the term. G.F. Munzel often discusses wisdom in relation to Kant’s pedagogy—cf Munzel 1999, 57ff and various passages throughout Munzel 2012. The reader may also consult the recent discussion at Ferrarin 2015, 80–103. There is an older and frustratingly truncated treatment—but a fascinating one—in Yovel 1980, 230–6.

¹⁰ This is in contrast to mere prudence (*Klugheit*), which directs action towards *happiness* (cf the footnote at GMS 4:416). For how prudence and wisdom relate the reader may refer to Engstrom 1997, especially 36ff. Engstrom argues that wisdom needs the *addition* of prudence, as practical savvy, to “complete” it. While I do not think this is correct for the critical Kant, the Kant of the 1760s seemed to hold exactly this view: see V-PP/Herder, AA 27:36.

¹¹ Fugate 2014 discusses this text in detail at 20–24.

¹² This passage, to all appearances, anticipates the more famous one at GMS, AA 4:391 (although wisdom is not directly mentioned there): “[T]he critique of pure practical reason [must] present the unity of practical with speculative reason in a common principle, since there can, in the end, be only one and the same reason, which must be distinguished merely in application.”

¹³ Kant insists on this throughout the corpus, e.g.: “The interest in [reason’s] speculative use consists in the *cognition* of the object up to the highest a priori principles; that of its practical use consists in the determination of the *will* with respect to the final and complete end” (KpV, AA 5:120).

¹⁴ See Kleingeld 1998, especially 322ff.

The world must be represented as having arisen out of an idea if it is to be in agreement with [moral] reason ... All research into nature is thereby directed toward the form of a system of ends, and becomes, in its fullest extension, physico-theology. (KrV, A815-6/B843-4)

Reason cannot *know* of any genuine moral teleology in nature, nor of any moral author. But the interests of reason allow it to *read* nature as *if* that teleology and author were present, for the sake of realizing its final end. In 1788 Kant puts it this way:

[S]ince a pure practical teleology, i.e., a morals, is destined to realize its ends in the world, it may not neglect their possibility in the world, both as regards the final causes given in it and the suitability of the supreme cause of the world to a whole of all ends as effect—hence natural teleology as well as the possibility of a nature in general, i.e., transcendental philosophy. This serves to secure objective reality to the doctrine of practically pure ends with respect to the possibility of the object in the exercise, namely the objective reality of the end that this doctrine prescribes as to be effectuated in the world. (ÜGTP, 8:182-3)

The discussions of reflective judgment and theology in the “Second Part” of the third *Critique* (KU, AA 5:359-484) carry forward and formalize this structure.

“Wisdom,” as a Kantian term of art, thus implies the coherence of the whole critical system: the unification of practical and theoretical reason by way of the teleology of the world, read as directed towards the final end of reason. This relies upon the *compatibility* of the idea of a moral world and the given order of nature, the same compatibility which serves as the spur in the “righteous Spinoza” version of the moral proof above. Albeit without mentioning wisdom,¹⁵ Kant says it is precisely “lawlike agreement ... with the ends [*eintreffende Zusammenstimmung zu dem Zwecke*]” (KU, AA 5:452)—the same language of “tuning together”—which the righteous Spinoza cannot currently expect from nature unless he elects for faith. And to tie the corpus together up to this point, Kant repeats exactly this same concept— with this same language—in a footnote on the second page of the “Failure” essay, stating that “the proper concept of wisdom represents only a will’s property of being in agreement [*zusammen zu stimme*] with the highest good as the *final end* of all things” (MpVT, AA 8:256).

When Kant says “the highest wisdom” in this first line, he is indeed still primarily talking about God. But it is a terminological shift that highlights the systematic and architectonic stakes of the whole essay. For the “Failure” does not deal only with the traditional topic of theodicy, but also specifically with the systematicity of Kant’s own thought.

2. Counterpurposiveness and the Fall of Righteous Spinoza Theodicy

If Kantian wisdom relies upon the ultimate compatibility between morality and the world of nature, then it is precisely that compatibility which the plaintiff of the “Failure” essay will attack. Kant’s second shift of terminology in that same opening sentence indicates how the attack is waged: like the shift from “God” to “the highest wisdom,” the substitution of “the counterpurposive [*zweckwidrig*]” for “evil” is a deliberate choice that directly implicates Kant’s own thought. Once again, the change in terms is not to suggest that the essay isn’t about the “problem of evil” but that it is so in a highly specific way. *Zweckwidrig* has no easy English translation: in everyday talk it means something like “inappropriate” or “counterproductive.”¹⁶ Kant here uses it quite literally in the sense of *against the end*, to refer to phenomena that are not just unfortunate or even malevolent but *radically contrary* to, even *incompatible* with, the final end of practical reason as such.¹⁷ Once the terms become clear, so do the stakes and the nature of the challenge—and also why the RST is helpless to meet it.

Kant presents the plaintiff’s case extremely generally, to their advantage: if he does systematically divide the objections the plaintiff makes against wisdom into three kinds,¹⁸ nevertheless their real challenge is simply insisting on the presence of the counterpurposive itself. As Kant writes it, the “charge which reason brings against [the highest wisdom] for whatever is counterpurposive in the world” never implies any overambitious

¹⁵ Wisdom *is* directly mentioned ten pages earlier in a related discussion on the limitations of physical teleology: “What help is it, one may rightly complain, to ground all [natural] arrangements on a great and for us immeasurable intelligence, and have it arrange this world in accordance with its intentions, if nature does not nor ever can tell us anything about the final aim, without which, however, we can form no common reference point for all these natural ends, no teleological principle sufficient for cognizing all the ends together in a single system as well as for forming a concept of the supreme intelligence, as the cause of such a nature ...? In that case I would, to be sure, have an **artistic intelligence**, for various ends, but no **wisdom**, for a final end, which, however, must really contain the determining ground of the former” (KU, AA 5:440-1).

¹⁶ Although it is used in passing in earlier texts, the first place where Kant deploys counterpurposiveness as a major term is in the third *Critique*, mainly in the discussion of the sublime as “contrapurposive for our power of judgment” (cf KU, AA 05:245-260). But in his discussion of physico-teleology he also uses this term to describe the theology of the ancient polytheists, who (he says) “found the good and evil in this world, what is purposive and what is contrapurposive, to be very mixed, at least as far as our insight goes, and they could not allow themselves to assume for the sake of the arbitrary idea of a most perfect author that there are nevertheless wise and beneficent ends lying hidden beneath this” (KU, AA 05:439). It’s a passage which suggests that counterpurposiveness for this earlier Kant can indeed be “evil,” but does not yet have the same degree of tension with “wise and beneficent ends” as it will a year later.

¹⁷ Other scholars have noted the special nature of the term “counterpurposive,” especially in its relation to the “highest good” and moral teleology, e.g. at Kravitz 2020 467 and Fonnenu 2023 15. So far as I’m aware, however, that relation has previously not itself been connected to the systematically central concept of wisdom, nor to how that concept develops throughout the essay.

¹⁸ Those concerning the “absolutely counterpurposive,” the “conditionally counterpurposive,” and the “disproportion between crimes and penalties in the world” (MpVT, AA 8:256-7). I omit the details of these arguments for the sake of space, and refer the reader to the seventh chapter of Huxford 2020 for detailed discussion.

speculative claims—that God is not moral, that there is no God, etc.—as these could be easily swatted down by appeal to the limits of theoretical reason. Rather, the charge made against the highest wisdom of the world really just amounts to insisting that phenomena exist within the creation of the highest wisdom (God) which are contrary to wisdom’s own final moral end. Put in more Kantian terms, the plaintiff’s complaint is that there is something in the order of nature that cannot coherently be made to fit moral teleology. It posits a break between the “*artistic wisdom* of the author of the world,” which we read as the cause of the natural order of the world, and the “*moral wisdom*,” which we read as directing that order towards the highest good as the final end of reason (MpVT, AA 8:256).¹⁹ This is a challenge analogous to the same tension found in the traditional theodical “problem of evil”—between an omnipotent, omnibenevolent God and the existence of evil—except that Kant has constructed a version of that challenge which aims for the heart of *his own system*. In these very first lines, his plaintiff challenges the compatibility between morality and nature which was supposed to make possible the final unity of theoretical and practical reason, and thereby the completeness of the system as Kant had previously envisioned it.²⁰

Because the objection to God is not speculative (unlike “Spinoza,” the plaintiff here posits no theoretical doubts about God) but moral and teleological, this same opposition between wisdom and the counterpurposive also seems to challenge the RST of just a year earlier. In the RST, faith in a moral creator could function as a resolution to the problem of evil because it meant that nature could be harmonized with moral ends, such that those ends could be realized within nature. But through suggesting the presence of phenomena in nature expressly discordant with such ends, the plaintiff now puts that harmony into question. It is as if the righteous but unbelieving “Spinoza,” a year after the third *Critique*, had fully appreciated the arguments implied by that passage and nevertheless turned them on their head. Kant had written then that if such a man “[assumes] the existence of a moral author of the world ... from a practical point of view” then he will find the final end of reason no longer “impossible”—yet Kant made that appeal while fully admitting the “deceit, violence, and envy” plus “poverty, illness, and untimely death” (KU, AA 5:452) which the moral agent sees in the world. The plaintiff need only point out that faith in a moral God cannot remove or justify such phenomena; if anything faith makes their presence still more egregious, since they are not just unfortunate but perhaps incomprehensible within a world stemming from the creator one is supposed to have faith in. And if a faithful moral agent can no longer find natural teleology to be morally trustworthy, then their projects seem to be “limited” once again: they may have as little ability to comprehend why their aims should be realizable as the righteous unbeliever.²¹

If I spend so much time insisting on two terms of Kant’s very first line, it is to highlight the stakes of this essay. The plaintiff’s challenge is posed not just against God, not just against the integrity and comprehensibility of Kant’s system (the unification of theoretical and practical reason which Kant had been seeking since the critical project’s first sketches), but also against moral faith itself. It thus might reasonably be expected that Kant should side with the “would-be advocates of God” as he had done in the *Lectures* of less than a decade earlier, and throw everything he has into a defense of wisdom and the compatibility of physico-teleology and ethicotheology. This would have been easy enough: insofar as the plaintiff stands as a *witness* to genuine counterpurposiveness, one only needs to dispute their witness and explain away all apparent evils as items which, given the right context, can still be fit into moral teleology (or which we at least can’t know to be absolutely evil).²² The following pages of the “Failure” instead do the opposite, ridiculing the very project of “doctrinal” theodicy and dismantling the arguments of its advocates. Kant shows contempt for the defenders

¹⁹ Kravitz 2021 suggests relating the different modes of wisdom in this footnote (artistic and moral) to the divine attributes listed on AA 8:257, so that this is equally a break between what Kant calls holiness and goodness. I read these distinctions as landing differently. In the footnote on AA 8:256 Kant insists that art—even elevated to a level beyond human reason, as artistic *wisdom*—remains directed only to “*optional ends*,” so that it can at most only be the wisdom of a demiurge unless it is united with morality. The “attributes of the world-author’s supreme wisdom” on the next page, by contrast, suggest *necessary* ends each intimately tied up with the highest good as discussed at KpV, AA 5:109ff: moral virtue (holiness), natural happiness (goodness), and proportionality between them (justice). They are thus aspects of a highest wisdom that is *already unified*, i.e. art and morality teleologically harmonized. Thus although cases of counterpurposiveness will relate to select attributes (and select defenses will be raised), theodical objections properly assault the teleological unity underlying all of these attributes as a whole.

²⁰ Here I agree with Christoph Schulte, who reads the concept of counterpurposiveness as a direct challenge to “the highest good” and thus a break from Kant’s earlier theodicy (Schulte 1991, 383–4)—while also suggesting that the stakes of this break extend to Kant’s vision for the entire critical system.

²¹ In principle, I think this same problem applies to psychologically-driven readings of the RST passage as well as to the “rational instability” readings I endorse. Whether it is a motivational boost or a rational satisfaction that faith in a moral God is supposed to supply, it does not supply those items without *context*. It might be possible, e.g., to get one’s moral resolve out of a pill or a shot and thereby render the problem of evil moot, but that’s not what Kant says. For Kant, at least, it is supposed to be because the righteous agent feels that they can *trust and rely upon* God to render their moral ends possible that they gain appropriate motivation or rational satisfaction. For both readings it is that trust which the plaintiff attacks, in that it becomes unclear how this God can be trusted to promote one’s end when the same God has allowed the existence of phenomena counter to that end. (Just so, one might still give up on the requirement that all of this be clear, which is where Kant is heading.)

²² Brachtendorf 2002 makes a version of this case on Kant’s behalf, that because we cannot *experience* counterpurposiveness in a strict sense it cannot be used as a challenge against divine wisdom. But I take such debates about whether we “really” experience evil, or whether it is “really” in the world, to miss Kant’s point. As we shall see, what makes the plaintiff (and Job) morally praiseworthy is that they are *sincere*, that they stand by their witness of evil despite incentives in the other direction. Indeed, that sincerity seems valuable enough to Kant that he is willing to revise the scope of his system to save it. Even if the plaintiff’s arguments are not *true* [wahr], their speaking them is *truthful* [wahrhaft], and this has its own worth. Defenders of divine justice like Job’s friends, Kant will say, “might convey more of an appearance of greater speculative reason and pious humility” than someone like Job, but theological ingenuity is no equal to sincerity: “Their malice in pretending to assert things into which they yet must admit they have no insight, and in simulating a conviction which they in fact do not have, contrasts with Job’s frankness ... much to his advantage” (MpVT, AA 8:266). Especially regarding moral evil, “the vindication is worse than the complaint” (MpVT, AA 8:258) because it is

of the highest wisdom and a certain sympathy for its challengers, in spite of how deeply they cut into his project. To be sure, this sympathy is not because he thinks they are any more knowledgeable than the other side: although there is no testing of the plaintiffs that runs parallel with the examination of the self-appointed divine defenders, Kant is nevertheless clear that “the doubts²³ raised against [wisdom] ... as objections, so far as our reason’s inherent insight regarding them goes ... [also cannot] prove the contrary” (MpVT, AA 8:263). I further explore Kant’s sympathies for the plaintiffs below, but even if he thinks they cannot prove much he also clearly believes they are *right* not to dismiss or explain away the counterpurposive: throughout his responses to the doctrinal theodicies, especially those concerning moral evil, Kant firmly rejects any account by which we could cease calling such phenomena evils. Thus, even if Kant cannot admit that some phenomena in the world really are discordant with the moral ends of reason, he also cannot deny it, and he will grant the plaintiffs every right to *feel* as they do. This seems to be a profound self-mate on Kant’s part: with the defeat of the doctrinal theodicies and the refusal to dismiss genuine counterpurposiveness, he appears to undermine the justification for rational faith in a moral author of the world and shatter all hope for systematicity.

3. Why Negative Wisdom?

It is the project of the difficult second half of Kant’s essay to develop what can be recovered of his project beyond the legitimacy of the plaintiffs’ challenge. If what is at stake is the compatibility between the order of nature and the highest good—and, as in the RST, the resolve and stability of the faithful agent who relies upon that compatibility—then it is *faith* which grants sight of its possibility. Thus the remainder of the essay amounts to a reinterpretation and reworking of what a Kantian life of faith should look like, given that we cannot dismiss counterpurposive phenomena. At the same time it is a drastic reconsideration of the concept of wisdom and its role. He is clear in the discussion to come that, whatever his sympathies for Job and similar plaintiffs, they still have the shortcoming of speaking “unwisely” (MpVT, AA 8:266) —implying that even if it must be revised (ultimately in the direction of apophaticism), the concept of wisdom need not be given up altogether. To set up a reading of the positive work of this essay, which is ultimately the development of a new, negative concept of wisdom by way of what Kant calls “authentic theodicy,” it will be useful to once again contrast this essay’s problem with that of the RST.

So far we have dealt with two distinct theodical protagonists from Kant, both moral agents but each with their own dilemma. The protagonist of the RST, the righteous Spinoza, is presented as a moral agent lacking in faith, or at least the right kind of faith. They do not see the matter of nature around them as counterpurposive (*zweckwidrig*), but “purposeless [*zwecklos*]” (KU, AA 5:452)—not explicitly counter to the final moral end they are committed to, but not amenable to it. Faith in the existence of a moral creator of the world is supposed to resolve this because it renders nature morally teleological, and thus makes their end—the realization of a moral world—really possible.

By contrast, even before Kant introduces Job the plaintiff of the “Failure” essay faces not a general lack of purpose in nature, but the presence of specific phenomena which run starkly against moral ends.²⁴ This is already a more complex situation: it would not even be sensible to pose such a complaint unless the plaintiff *already expected* a general moral teleology in nature and thus a consistent “*Zusammenstimmung*” with the final ends of reason, something which (by Kant’s own arguments, above all the RST) they could not presume unless they already had faith in a wise moral creator. So the plaintiff’s problem, it seems, is not a lack of moral faith, but that their faith is *troubled*. Their faith first grants them the capacity to interpret the world as morally teleological, yet that very interpretation then reveals items which seemingly cannot be reconciled with moral ends. Rather than a purposeless world, they confront an *otherwise* purposeful world whose evils are incompatible with the God whom they rely upon. Thus any project to resolve the troubles of this plaintiff’s faith faces a dilemma with at least three horns. (1) If the plaintiff gives up on faith altogether, they return to the same limitations as the righteous Spinoza above. (2) If they accept what Kant calls “the system of *unconditional divine decision*” (MpVT, AA 8:265), where God does whatever God wants (moral or not), the world can remain purposive on pain of being morally arbitrary. (3) If they attempt to explain away the counterpurposive through the doctrinal theodicies that Kant is at pains to examine, then they do so at the risk of insincerity (discussed below).

This dilemma, novel in Kant’s work, cuts out his work for him: if the moral agent can neither renounce faith in the wisdom of a moral creator nor deny the presence of the counterpurposive in creation, it remains to be

insincere, and insincerity—even for theological ends—is morally perilous. For the purposes of this text, the ontology of evil is thus not as important as the moral agent’s witness of it, and reaction to it.

²³ The English word “doubt,” although a correct translation, may be too strong for the “Failure” essay’s *Zweifel*. “Doubt” often has the tone of a fully worked out, fully endorsed epistemic stance (namely that something cannot legitimately be held as true). The “doubts” of the Kantian plaintiff are more like qualms or misgivings against the highest wisdom than explicit “dubitantes” in this stronger sense.

²⁴ It is crucial above all not to read the plaintiff as founding their case upon *their own* suffering, even if undeserved. Kant is very clear in a footnote that the morally righteous should not expect any guarantee of happiness from divine justice: “For under divine rule even the best of human beings cannot found his wish to fare well on divine justice but must found it on God’s goodness [*Güte*], for one who only does what he should [*seine Schuldigkeit*] can have no rightful claim on God’s doing him well [*Wohlthun Gottes*]” (MpVT, AA 8:258, translation amended). Concern over divine justice, Kant insists here, “is directed not at the well-being which does not befall the good, but at the ill which does not befall the evil.” Thus the objections of Kant’s Job and similar plaintiffs should not be that their personal righteousness is answered with suffering rather than divine rewards, because they have no right to such rewards. Those complaints rather have to be taken up into a broader challenge against that very “goodness [*Güte*] of the legislator” as “directed to the well-being of his subjects” (MpVT, AA 8:257-8), in its disharmony with the apparent evils in creation. On this point, especially as regards Job, see Kravitz 2021, 352-58.

seen whether the plaintiff's challenge can still be addressed in some way that leaves these tensions intact. What Kant needs to "answer" the plaintiff, in other words, is some healthy way to accommodate both sincerity in the face of evil and moral faith within the living comportment of the moral agent. It is this possibility that Kant presents under the name "negative wisdom," which serves as less a resolution of a "problem of evil" than a circumvention of the forces that made it a problem.

Kant presents and develops the concept of negative wisdom in a series of steps in the middle of the essay, which I will address in sequence. Two paragraphs at the center of the text introduce the term in a formal and initially unsatisfying way, so that it will be useful to explore some reasonable objections as a lead into the account of "authentic theodicy" that follows. Authentic theodicy and Kant's accompanying reading of Job can be best understood, I argue, as filling in the details of what negative wisdom is and what it is supposed to accomplish. The passages involved are quite dense and difficult, and I will supplement Kant's word by thinking through the "phenomenology" of the concepts at issue.

4. Negative Wisdom I: Initial Presentation and Objections

This new kind of wisdom is introduced directly after Kant's discussion of the doctrinal theodicies.²⁵ In order to put an end to the trial between the plaintiff and the self-appointed defender, Kant says, we can and must reach a "negative wisdom," here described as "insight into the necessary limitation of what we may presume with respect to that which is too high for us" (MpVT, AA 8:263). The possibility of a reconciliation between artistic and moral wisdom is salvaged by simply doing away with the *comprehensibility* of that reconciliation. Kant now accepts that the unity between the kinds of concepts we must "think together in the idea of a world and of a highest good" (MpVT, AA 8:264) not only cannot be encountered in experience, but is also entirely beyond the concepts of both speculative and practical reason. The trial must be brought to an end, therefore, because we lack any concept "of the *unity in the agreement* in a sensible world between ... artistic and moral wisdom ... nor can we ever hope to attain one" (MpVT, AA 3:263). And just because we cannot grasp how a partly counterpurposive world can adequately serve the purpose of a moral creator, Kant suggests, we should not conclude that moral and artistic wisdom cannot ever be unified—it simply means that we have no insight into how it would be possible.

This first account of negative wisdom is, first of all, a serious retreat in Kant's conception of how far human reason can cognize any "highest wisdom," as well as in his ambitions for a teleological unity of reason. And insofar as it reads on its face only as a flat denial of our ability to know anything concerning the divine, even a charitable reader of Kant has right to be concerned that this Kantian "negative theology" fails even by Kant's own lights. In this initial account Kant can appear to be cheating some of his own rules and principles in at least two ways, and it is worth exploring such concerns:

1. This same essay earlier told us that the theodacist may not simply make a "decree of incompetency of the tribunal of reason" and thereby "dismiss [*abweisen*] the complaints with a concession of the supreme wisdom of the author of the world" (MpVT, AA 8:255). We were promised that theodicy could not rule the court of reason simply unqualified to deal with the question, and yet it appears that Kant is doing exactly that. It is gross inconsistency for Kant to reject this tactic there and deploy it himself eight pages later.
2. Perhaps even more dire from a Kantian point of view, the argument here is *unsatisfying*. Elsewhere in the corpus Kant will insist that reason's falling into error and conflict with itself is always the result of certain "positive cognitions that belong in the domain of pure reason" (KrV, A795/B823)—cognitions which, viewed with a sufficient critical eye, can still have other legitimate roles. Both parties in this theodical dispute, the plaintiff and divine defender, seem to represent valid rational interests (sincerely calling the counterpurposive what it is versus insisting on the moral wisdom of creation), however opposed they may be in the details. So by Kant's lights we should expect to see such interests given some meaningful satisfaction—perhaps not in the form the parties expect, but still some resolution to their rational needs. Yet Kant's "proof" of the incapacity of our theological insight seems to do nothing of the sort: the plaintiff seems left with their faith just as troubled as ever, while the theodacist is left devoid of concepts with which to justify such faith.

The second concern in particular shows why Kant could not end the discussion just by showing the lack of any adequate route to knowledge. Even if Kant thinks he can prove that a negative wisdom is "within our reach" (MpVT, AA 8:264), whatever way he "dismisses" theodical debates must still satisfy the interests of reason at issue. I suggest that Kant responds to such concerns though developing his concept of negative wisdom in a second step, by way of "authentic theodicy" and his accompanying commentary on Job.

5. Negative Wisdom II: Authentic Theodicy and Job

Negative wisdom is "insight" into our theological limitations, but as negative *wisdom* it must also align us—tune us together with all of our capacities in harmony—towards the final end of practical reason. In that regard, it must represent a way to maintain firmness of moral faith without explaining away counterpurposive evil. The account of authentic theodicy that follows begins to show how this is possible, answering the concerns above in the process.

²⁵ Kant explicitly mentions "negative wisdom" once again seven years later, at SF, AA 7:93. Once again it is a wisdom specific to human beings (and contrasted to divine providence), but in this case the wisdom regards the gradual reduction of war.

Before his commentary on Job (which is his real illustration of authentic theodicy), Kant presents another formal characterization. His choice of terms defuses the first concern above, and also clarifies the project. Kant characterizes authentic theodicy as one kind of “*interpretation*” of nature insofar as God announces his will through it” (MpVT, AA 8:264), just as with the doctrinal theodicy (“theology proper,” the theodicy of defending God’s cause) which Kant dealt with in the first half of the essay. But whereas doctrinal theodicy deals in the “rational inferences” of a merely “*ratiocinating* [vernünftelnden] (speculative) reason,”²⁶ the authentic one is supposed to be a “decree” (*Machtspruch*) made “by the law-giver himself”—and for Kant it makes no difference whether this is the moral God as such or an “*efficacious* [machthabend] practical reason,” the kind of reason “through which we form our concept of God ... as a moral and wise being.”²⁷ Kant calls this authentic kind of theodicy the “mere *Abfertigung* of all objections against divine wisdom, if this dismissal is a divine decree” (MpVT, AA 8:264). Now the Cambridge translation is misleading here: it renders *abfertigen* as “dismiss,” the same word used to translate *abweisen* earlier. So the first concern above sees a contradiction here—that Kant writes in the first pages that the author of theodicy is not allowed to “dismiss” [*abweisen*] theodical complaints (MpVT, AA 8:255) but here seems to do that himself—when there really isn’t one. *Abfertigen*, the term Kant deploys in the authentic theodicy passage, can indeed shade into some of the same meanings as *abweisen* (e.g., as “disposing” of something) but it just as commonly means “to process” or “to handle,” e.g. as in making sure a piece of mail ends up at the right place.²⁸

What authentic theodicy is supposed to be doing, then, is “handling” theodical objections in the right way, something which can only be done (Kant says) by the direct word of the moral God or by efficacious practical reason. And this handling cannot be “dismissive,” i.e. it has to *satisfy* the interests of reason noted in the second concern from above. The climax of the Job story, for Kant, is clearly supposed to represent a case of authentic theodicy “in action,” and it is here that Kant develops the previously vague negative theological language of 8:263-4 (which presented negative wisdom only as our inability to know) into something more substantive. As if he had been taking notes from John Scottus Eriugena, Kant presents the inscrutability of the creation—and of the highest wisdom behind it—as a *privilege* of creation, as exactly what must be seen and appreciated. It is the sight of this which renders Job a wiser man.

Kant reads the central theme of the climax of Job to be *divine wisdom as given within creation*, or more exactly the conviction of the faithful moral agent (Job) into it.²⁹ Kant does not locate the authentic theodicy in the direct speeches of God from Job 38-41—he does not quote these passages—but simply writes that at the end of the story “God deigned to lay before Job’s eyes the wisdom of his creation” (MpVT, AA 8:266). The authentic theodicy Kant has in mind, that is to say, is not a particular line of reasoning or even a spoken decree, but the *revelation* of wisdom as given through the world. The passage that follows is worth quoting in full:

He allowed [Job] glimpses into the beautiful side of creation, where ends comprehensible to the human being bring the wisdom and the benevolent providence of the author of the world unambiguously to light; but also, by contrast, into the horrible side, by calling out to him the products of his might, among which also harmful and fearsome things, each of which appears indeed to be purposively arranged for its own sake and that of its species, yet, with respect to other things and to human beings themselves, as destructive, counterpurposive, and incompatible with a universal plan established with goodness and wisdom. And yet God thereby demonstrates an order and a maintenance of the whole

²⁶ On the relation of *vernünfteln* or “rationalizing” to moral self-deception, see Noller 2022. Interestingly Kant uses this same term further down to describe the arguments of both Job’s friends and Job himself (“...was beide Theile vernünfteln oder übervernünfteln...” (MpVT, AA 8:265), implying that Job too—in all his sincerity—is not immune from excessive rationalizing.

²⁷ This equation of the direct word of God with the pronouncement of practical reason is, of course, crucial for Kant’s employment of Job: it gives him the right, as it were, to read this biblical story as presenting concepts directly relevant to practical philosophy.

²⁸ Throughout the essay Kant poses outright dismissal (*Abweisung*) as one mode of “handling” challenges to the highest wisdom, but not the only one. This is clearer in the context of the full opening consideration of dismissal at MpVT, AA 8:255. The author of a theodicy, Kant first says, “is not therefore allowed to dismiss [*abweisen*] the complaints in the course of the process of law through a decree of incompetency of the tribunal of human reason (*exceptio fori*).” Then he continues, explaining the details of what this dismissal consists in: “... i.e. he cannot handle [*abfertigen*] the complaints with a concession of the supreme wisdom of the author of the world, imposed upon the plaintiff, which would immediately explain away as groundless even without examination, all doubts that might be raised against it” (translation amended). This is to say that *Abweisung* is one kind of *Abfertigung*, one that explains away without examination. And this is not always wrong: indeed, one of Kant’s stated ambitions for negative wisdom is to reach a point where “all further attempts by a putative human wisdom to gain insight into the ways of the divine wisdom are fully dismissed [*abgewiesen*]” (MpVT, AA 8:263). So dismissal is the right mode of “handling” overambitious claims to insight, but not legitimate doubts against divine wisdom.

²⁹ It is well beyond my project to discuss Kant’s approach to biblical interpretation more generally (see Lema-Hincapié 2006, Palmquist 1992, and Pasternack 2015 for good starting points to a broad topic), but his use of Job here deserves special comment towards that question. He will tell us in the *Religion* that “since... the moral improvement of human beings—constitutes the true end of all religion of reason, it will also contain the supreme principle of all scriptural exegesis” (RGV, AA 6:112), and in the *Conflict of the Faculties* that “if [a scriptural text] contains statements that contradict practical reason, it *must* be interpreted in the interests of practical reason” (SF, AA 7:38). Now such rules, read on their face, can suggest that Kantian exegesis must flatten all scripture into straightforward moral instruction, or at least prevent all tension with moral interests. It can seem that such texts can then teach nothing novel beyond what we could learn by (e.g.) reading Kant scholarship. Yet at moments in the corpus where Kant actually engages with scripture at length—here with the reading of Job, with the book of Revelation at EAD, AA 8:333ff, arguably also with Genesis at MAM, AA 8:110—his readings, and the lessons he suggests, are often far more paradoxical than that first impression might suggest. Given his own interpretive strictures, there is seemingly no reason why Kant could not have dismissed the whole scenario of Job—of God choosing to bring a virtuous man to ruin—as morally abhorrent, as he suggests with the sacrifice of Isaac in Genesis 22 (cf RGV, AA 6:87 and 6:187). Instead, as below, he uses the climax of Job to develop a perplexing and counterintuitive consequence of his own thought. This suggests that even if the conclusions must still remain in line with practical instruction, Kantian exegesis can very well be used to power quite nuanced and novel insights.

which proclaim a wise creator, even though his ways, inscrutable to us, must at the same time remain hidden—indeed already in the physical order of things, and how much more in the connection of the latter with the moral order (which is all the more impenetrable to our reason). (MpVT, AA 8:266)

Kant's Job beholds beautiful entities, *and* counterpurposive entities, *and* their interrelationships within the same holistic natural order—yet Kant insists in all of this that what is revealed above all is the creation itself, “especially its inscrutability [*vornehmlich von Seiten ihrer Unerforschlichkeit*].”

If Kant earlier insisted that negative wisdom was something that we could and should reach, Job is clearly supposed to exemplify what reaching such an insight is supposed to look like, so that this biblical interpretation is supposed to put flesh on the formalistic bones we previously received. Creation and the ways of divine wisdom are put directly before Job's eyes as “inscrutable,” “impenetrable,” “hidden,” and the sense of these passages turns on how to interpret this “hiddenness” of wisdom and creation. The appeal to inscrutability, for example, can easily be read as craven self-contradiction on Kant's part, as the full embrace of at least one of theodicies that Kant himself had already rejected. If Kant means to suggest that these teleologically mysterious ways of God may become known to us at some point in a future world, then he relies upon the same speculative promises found in theodicies II.c and III.c—which were previously rejected as “pretended” (MpVT, AA 8:260) and “arbitrary” (MpVT, AA8:262). Still worse, if the operations of divine wisdom are hidden because they are alien to human reason, then this appears to collapse into the God of alternative ethics from theodicy I.a—a view which, Kant said, must be “given over to the detestation of every human being who has the least feeling for morality” (MpVT, AA 8:258).³⁰ But this is to project speculation onto Kant which his prose seems at pains to avoid. The “inscrutability” he insists upon does not refer to anything held back from what Job sees (whether future explanation or alternative moral principle), but *is itself* the very item that the authentic theodicy reveals. Kant is suggesting that divine wisdom and the creation are properly grasped when grasped as *hidden*.³¹ There is no additional insight to expect; the insight *just is* active witness to the inscrutability.

As an “answer” to the problem of evil this may not be what the reader would have expected, and Kant even calls the authentic theodicy only a “vexing resolution” of [Job's] doubts (MpVT, AA 8:267). The resolution is vexing [*befremdliche*] because we expect to find a concept that can somehow release the tension between wisdom and the counterpurposive that has stood since the first line of this essay. But there is just no concept to be had: the counterpurposive within creation is not justified or explained away, but simply allowed to stand as exactly the phenomenon “incompatible with a universal plan established with goodness and wisdom” (MpVT, AA8:266) which the original theodical plaintiff had always seen it as.³² What Job reaches is what Kant calls “*bloß die Überführung von seiner Unwissenheit*,” (MpVT, AA 8:267) merely the conviction of his un-knowing,³³ and Kant has in mind not just awareness of his lack of knowledge but a “transfer” (*überführen*) into a new and principled stance. Plaintiffs such as Job have always been right in refusing to ignore the counterpurposive and its tension with the final end of reason—indeed, if what they challenge is only the *comprehensibility* of wisdom and creation then they have spoken rightly, as Kant now gives up such comprehensibility. The only item which such plaintiffs have lacked—why Job initially spoke “unwisely”—is the “vexing” appreciation that inscrutability is a *positive and ineliminable characteristic* of divine wisdom and creation. Kantian negative wisdom is, first of all, just this hard-won appreciation—not just intellectually accepted as the lack of an answer but (to anticipate the next sections) lived out.

That is certainly not all it must be, however. Huxford is correct to write that at this point “Kant is no longer offering any explanation” (Huxford 2020, 132) for counterpurposiveness, which means that the plaintiff from the beginning of the essay is still left surrounded by the same counterpurposive phenomena that gave rise to their doubts in the first place. If negative wisdom cannot vindicate divine (or Kantian) wisdom against doubts posed by the counterpurposive, if it cannot even justify moral faith as a way to render moral agent's aims comprehensible (as in the RST), still Kant's “vexing resolution” must actually satisfy the rational interests

³⁰ Huxford 2020, 130-2 presents more fully developed versions of such objections.

³¹ Although not immediately relevant to his argument, it is notable just how close Kant comes to the language of a thinker like Eriugena in these passages. That God is inaccessible to knowledge is a common point of theological orthodoxy, but for Eriugena—as for Kant—God's “darkness” and “hiddenness” represent not just shortcomings of the human intellect but a feature of the highest entity's radical eminence. “[T]he Divine Essence is comprehensible to no intellectual creature” (447c at Eriugena 1987, 31), Eriugena says—and this includes the superhuman (angels, etc.) just as much as us—because it is *greater* to be inscrutable: “[S]o great is the splendour of the Divine Goodness that, not unreasonably for those who desire to contemplate it and cannot, it shall be turned into darkness. For He alone, as the Apostle says, ‘possesseth the inaccessible light’” (681b at Eriugena 1987, 308). And if one of the key moves in Kant's present text is to transpose this same hiddenness to nature, much the same move is to be found in Eriugena as well: “For just as God as He is in Himself beyond every creature is comprehended by no intellect, so is He equally incomprehensible when considered in the innermost depths of the creature which was made by Him and which exists in Him” (443b at Eriugena 1987, 27). To be sure, the kinship in language and concepts here does not seem to me to place Kant comfortably in the tradition of “negative theology”: where that tradition will typically take this incomprehensibility as a route towards mysticism and unification with the divine, Kant instead drives it back towards morality and “good life conduct.” It is unhelpful to strain the comparison and I would be surprised if Kant had ever read a word of the *Periphyseon*, but it is still useful to point out the kinship as a tool to interpret the strategy of the “Failure.” For a useful introduction to Eriugena I point the reader to Otten & Allen 2014.

³² Although in a sense it is correct to say that the counterpurposive has a “role” here—it is, after all, the presence of the counterpurposive within the whole which so starkly *shows* the inscrutability of creation—it is important not to suggest that the counterpurposive somehow exists *because* of that role, as if moral evil and the rest of the world's ills had been created in order to reveal the hiddenness of divinity to the faithful. Kant would reject wrapping evil up into such a tidy story, and rightly so. The counterpurposive is just that, something *wrong* in the order of the world, and it will always remain so—regardless of how far the moral agent personally makes peace with that order.

³³ The Cambridge translation renders this as “merely from being convicted of ignorance,” which bizarrely renders the key noun into a predicative expression.

posed in the second concern above and offer reason something beyond an empty quietism. Indeed, I argue that there is a genuine satisfaction to be had for the rational interests of both the plaintiff and the theodicist. Despite the admittedly terse, dense, and nebulous character of Kant's prose in these passages, it's pretty clear that the "conviction" of negative wisdom—what authentic theodicy is supposed to accomplish—consists in some kind of *personal stance* towards the situation of creation and divine wisdom. What I suggest below is that this stance also concerns where and how the moral agent stands within that creation: negative wisdom is supposed to be as much a *self-interpretation* in the context of "good life conduct" (MpVT, AA8:267) which, as lived out, accompanies the theological side of negative wisdom already discussed. If the way that authentic theodicy "handles" doubts does not *eliminate* doubts, it may nevertheless change the relation that the faithful moral agent has towards themselves and, by extension, how they interpret such doubts.

Making this case requires thinking through the "stance" of the faithful moral agent in a concrete way, for which I will step beyond Kant's word and into a freer exploration of the concepts involved—first concerning faith and the counterpurposive, and then concerning sincerity.

6. The Phenomenology³⁴ of Counterpurposiveness

Faith and counterpurposiveness, when lived out, have a distinctive phenomenological dynamic. For Kant, moral faith [*Glaube*] in the existence of God is a kind of accompaniment to morality, insofar as the agent wills the realization of a moral world. If their willing is not to be in vain, then the moral agent must presuppose conditions which make such an end possible, and (for Kant) that includes the assumption that the creator of the present world is themselves moral and intends the same ends. It is this practical faith in a moral creator that allows the agent to then read this world as morally purposive, i.e. as suitable for the realization of a better one. And so far, so good; this is the foundation upon which the RST is built. But the phenomenological manifestation of the world as *morally purposive* is itself the condition for the manifestation of the *counterpurposive*, exactly what the plaintiff objects to in divine wisdom.³⁵ This is crucial: a character "entirely indifferent in regard to moral questions" (KrV, A829/B857), and therefore unaffected by the RST or any other moral proof of God, could very well take note of many moments of misfortune or unfairness, or many vicious people—but they could never see them as *counterpurposive*, only as unfortunate natural facts. The only character who would be able to witness such phenomena as *genuinely* counterpurposive, i.e. as contrary to the final end of reason, is a moral agent committed to the realization of that end, and thereby faithful in a morally wise author of the world who is supposed to make that end possible.³⁶

This dynamic further explains Kant's earlier sympathy for the plaintiffs, as well as his suspicion of traditional theodicy as "defending God's cause." Moral faith and the phenomenological visibility of the counterpurposive should have a direct relationship: the more an agent wills a better world, and the more they hold to their faith, the more evident and egregious the evils of the present one will become. That suggests, first, that sincerely expressed doubts and complaints represent not a rejection of moral faith but if anything a confirmation of it, as those who are the most morally faithful should be the most aggrieved by the evils of the created world. Job, the character who refuses to deny that evils really are evils, clearly stands as more morally faithful than the interlocutors who do. This further suggests that the project of doctrinal theodicy, insofar as it seeks to resolve or remove such doubts, not only lacks the means to "defend God's cause" but may actually *harm* that cause in some cases. Kant's concerns about sincerity at the end of this essay, specifically concerning oath-taking as a mode of "extorting truthfulness" (MpVT, AA8:268), directly reflect this.

³⁴ Often in contemporary anglophone philosophy "phenomenology" is used to mean something like *the description of personal, empirical experience*, in a fashion not far removed from Brentano's "descriptive psychology." Here I mean more an account of the *internal interplay of the structures* which Kant describes in the abstract, so that I loosely follow Husserl's 1903 edict that "[phenomenology's] descriptions do not concern lived experiences, or classes thereof, of empirical persons; for of persons – of myself and of others, of lived experiences which are 'mine' and 'thine'—it knows nothing, assumes nothing. Concerning such matters it poses no questions, attempts no definitions, makes no hypotheses. In phenomenological description one views that which, in the strongest of senses, is given, just as it is in itself" (Husserl 1994, 251). For Kant, of course, it is precisely the comportment of persons (myself and others) that ultimately matters. But these structures clearly have an important interrelationship, and description of their dynamics will show how they matter to such comportment.

³⁵ Pasternack identifies this dynamic, and uses it to argue that there is something incoherent about how the "problem of evil" is usually posed: "Consider... that for Kant, the judgment that there is evil in the world as well as the belief in God are both matters of practical reason. They arise from the same faculty, and given that reason cannot come into conflict with itself, we may a priori deny [that they are incompatible]" (Pasternack 2023, 154). Doubts and misgivings must be taken seriously, to be sure, but they do not represent *incompatibility* as such. I ultimately disagree with Pasternack's contention that the fact of evil acts as a *reason* for faith—that "we come to faith because otherwise our encounter with evil compromises our moral resolve" (Pasternack 2023 157)—as I view this as counter to the concept of a faith founded on morality, but he is right to notice the connection between moral resolve and the experience of evil.

³⁶ To be sure, this dynamic opens questions which I don't think Kant's text can answer. If the faithful moral agent indeed sees the counterpurposive as counterpurposive and the morally indifferent atheist cannot (because they do not presume any moral purpose for the world to begin with), still that leaves a great deal of room for middle cases. Does the righteous Spinoza, for example, experience "evil" in this sense? It seems odd that, moral as they are, they would take no notice of gross iniquity at all, or even see it as merely unfortunate. Alternatively, consider a megalomaniac who interprets the world as teleologically directed only towards their own glorification: if they encounter phenomena "counter" to this end, do they encounter them as counterpurposive? Such questions remain valid and interesting even if ultimately they may not matter much to Kant (whose final concern seems to be that faith founded on morality can remain coherent).

7. Sincerity and Critique

Kant's lengthy concluding comment can have the appearance of a mere appendix, but especially when read with the dynamic above in mind—and also Kant's misgivings about doctrinal theodicy—it has a direct relation to the concept of negative wisdom.

Kant introduces sincerity [*Aufrichtigkeit*] by relating it to faith, even calling it “the principal requirement in matters of faith” (MpVT, AA 8:267). Sincerity (or “truthfulness”), he says, amounts to “formal conscientiousness,” the consciousness of having been cautious not to speak wrongly in a given case (MpVT, AA 8:268). It is a kind of *general interest* in only speaking truth: on occasion one may not get things right, but one can still have “immediate consciousness” (MpVT, AA 8:267) that one at least tried. This insistence on truthfulness goes beyond the usual duty that one not lie: the sincere person must actively scrutinize themselves not only about whether what they propose to say is really true but also about whether they themselves take faith in it. “[F]ormal conscientiousness,” Kant writes, “consists precisely in the care in becoming conscious of [our] faith (or non-faith) and not pretending to hold anything as true we are not conscious of holding as true” (MpVT, AA 8:268, translation amended). What Kant is concerned about here is not just lying (i.e. intentionally misstating the truth) as one moral failing among others, but specifically the misstatement of truth by way of *apathetic self-ignorance*, where “someone says to himself ... that he believes, without perhaps casting even a single glimpse into himself—whether he is in fact conscious of thus holding a truth or at least of holding it to some degree” (MpVT, AA 8:268). A lack of sincerity, for Kant, means not only that the agent might deceive others, but that they might cease to care about truth *at all*, even about what they themselves hold for true. And if agents are encouraged to thoughtlessly propound whatever is accepted and advantageous, there is no reason why they should notice if they do not actually maintain the faith they claim to maintain—and, accordingly, do not actually will so morally as they would suppose.

If these passages obviously anticipate Kant's *Religion* to come, this concept of sincerity also returns the “Failure” essay to the fundamentals of critical method.³⁷ If we cannot presume that we actually hold for true what we passingly suppose that we hold for true—i.e. if we must check up on ourselves—then the comportment of sincerity amounts to the search for, and insistence upon, *self-awareness* in regard to faith. In 1781, Kant described the call for a “critique of pure reason” in much the same terms:

[R]eason should take on anew the most difficult of all its tasks, namely, that of self-cognition, and ... institute a court of justice, by which reason may secure its rightful claims while dismissing all its groundless pretensions ... and this court is none other than the critique of pure reason itself. (KrV, Axi-xii; translation amended)

Self-cognition in that case concerned working out the capacities of reason overall, while sincerity in the “Failure” only concerns examining my own faith. But this is still a question of self-cognition, so that critique and sincerity share analogous functions: my sincerity is not itself my faith (and still less my morality), no more than critique is itself a system of pure reason (metaphysics), but in each case the one acts as a consistent check and corrective upon the other.³⁸ This is presumably the sense in which sincerity is a “principal requirement” for faith: even if it is admirable in itself, it also serves as a crucial *check* upon morality and faith overall. It is the lived concern that we are actually as faithful, and as moral, as we think we are.³⁹

The importance of this “check” is first evident in cases where it breaks down, and doctrinal theodicy—exemplified by Job's friends—can represent just such a case. Kant is well aware that under external pressures moral agents can “feign conviction... even in their inner profession,” an inner dishonesty that “can also have external harmful consequences (for it gradually forges actual persuasion).” The danger is that the faithful moral agent bombarded by doubts, if they are persuaded (through reasoning or outright coercion) that their

³⁷ But also back to the “natural dialectic” of the *Groundwork*. Indeed, seven years before his first essay on religion, Kant already insisted that the “propensity to rationalize against those strict laws of duty, and to cast doubt on their validity, or at least their purity and strictness and, where possible, to make them better suited to our wishes and inclinations, i.e. fundamentally to corrupt them and deprive them of their entire dignity” (GMS, AA 4:405), necessitates the correction of popular morality by moral metaphysics and critique (see Sticker 2017). This tendency towards anti-sincerity and self-deception, what Green calls “intentional suppression and distortion of information to maintain a certain belief, satisfy a certain desire, and maintain a certain emotional state” (Green 1992, 159), means less that agents will to disobey the moral law than that they intentionally refuse to think about their willing (and so about themselves) in any way that would create trouble for them. Even this early, then, Kant already had an account that places insincerity at the foundation of moral evil and suggests a mode of self-cognition as its countermeasure. See readings of this passage in Callanan 2019, 14–19 and Noller 2022, 180–82; more general discussions of Kantian self-deception can be found at Green 1992, 159–69 and Papish 2018, 67ff.

³⁸ The relation between sincerity and critique can also be confirmed in that both are meant to resolve parallel kinds of *illusion*, namely moral self-deception and speculative error. See Noller 2021.

³⁹ Crucially, however, in neither case is the fundamental tendency of reason to go awry—whether through transcendental illusion or moral self-deception—ever really removed by such a “check.” An anonymous reviewer on this paper correctly remarked to me that sincerity, as a comportment of the moral individual, is always carried out against the background of a tendency towards *insincerity* which is more fundamental, what Kant calls the “dishonesty, by which we throw dust in our own eyes and which hinders the establishment in us of a genuine moral disposition” (RGV, AA 6:38). In his account of radical evil penned the next year, Kant writes that “We cannot start out in the ethical training... with an innocence which is natural to us but must rather begin from the presupposition of a depravity of our power of choice in adopting maxims contrary to the original ethical predisposition; and, since the propensity to this [depravity] is inextirpable, with unremitting counteraction against it” (RGV, AA 6:51). But because moral self-deception is “inextirpable,” the work of such counteraction—including sincerity—is never complete. This, not only because the heart can never be fully purified, but because even our best behavior is perpetually suspect (and on this point, see Welsch 2019 especially at 69ff). At the extreme, even the “check” itself—insofar as we posture ourselves as sincere—is not immune from scrutiny, as it too may be insincere.

inner convictions must be incorrect, numb their own moral sensitivity to the point of “[distorting] even inner declarations before [their] own conscience” (MpVT, AA 8:268). If they learn to distort their own witness to the counterpurposive, then they lie about their own faith—a mode of lying which Kant calls “the most sinful, for it undermines the ground of every virtuous intention” (MpVT, AA8:269). Thus it is in this concluding comment that we get the best explanation for why Kant thinks that the doctrinal theodicies are not just fruitless, but can be (like Job’s friends) full of “malice” and “false flattery” (MpVT, AA8:265-6)—they are *insincere*, so that theodicy may actually serve to *lessen* faith and moral conviction rather than bolstering them.⁴⁰ When Kant makes the cryptic remark that “the vindication is worse than the complaint” (MpVT, AA8:258) in reference to theodicy I.a (the dismissal of moral evil), he may speak less concerning the deficiency of that argument than the degree of insincerity required for a moral agent to accept it.

In addition to that, however, the role of sincerity also turns out to be the foundation of negative wisdom. In the passage directly following Kant’s commentary on God’s revelation to Job (i.e. the authentic theodicy), Kant writes:

The conclusion is this: Since Job admits having hastily spoken about things which are too high for him and which he does not understand—not as if wantonly, for he is conscious of his honesty, but only unwisely—God finds against his friends, for (as conscientiousness goes) they have not spoken as well of God as God’s servant Job... Hence only sincerity of heart and not distinction of insight; honesty in openly admitting one’s doubts; repugnance to pretending conviction where one feels none, especially before God (where this trick is pointless enough)—these are the attributes which, in the person of Job, have decided the preeminence of the honest man over the religious flatterer in the divine verdict. (MpVT, AA 8:266-7)

Job speaks “unwisely” prior to God’s moment of revelation, but he is *sincere* (i.e. aware of, and insistent upon, his faith and his doubts) from the beginning. Sincerity, as a mode of self-cognition concerning faith, is therefore a prerequisite for the negative wisdom that the essay is concerned with.⁴¹ As a last thought, I will suggest that negative wisdom is in fact the *completion* of such sincerity—that it does not and cannot eliminate doubts, but it can give the moral agent the self-understanding and the context with which to carry on with their projects.

8. Negative Wisdom and Reflective Faith

At any point in his critical career Kant could have written, as in the pages here, that “[t]heodicy... does not have as much to do with a task in the interest of science as, rather, with a matter of faith” (MpVT, AA 8:267). But by the point of the “Failure” no theodicy, not even the authentic, can or should “answer” the problem of evil in the sense of removing faithful doubts against divine wisdom. There is a sense in which Huxford is correct to say that authentic theodicy “fails” as a defense of the highest wisdom of the creator (Huxford 2020, 137), but at this point Kant is also quite open that he is no longer attempting what he considers “proper” theodicy at all (MpVT, AA 8:264). Witness to the counterpurposive is an *aspect* of moral faith, not a confutation of it, so that it is insincere to deny such witness or explain it away. Authentic theodicy, as in Job, must let these tensions stand and even insist on their visibility, allowing the moral wisdom of the “beautiful side” of the world to hang in inscrutable unity with the counterpurposive elements within the same world.

But this “improper” theodicy is still supposed to be a resolution of doubts, and thus a satisfaction of rational interests sufficient for faithful moral agents to persist in their moral projects. Yet Kant’s last comments on Job, although they give us his best hint at what authentic theodicy is supposed to accomplish at a practical and rational level, are still by no means clear:

The faith, however, which sprang in [Job] for such a vexing resolution of his doubts—namely merely the conviction of his un-knowing—could only arise in the soul of a man who, in the midst of his strongest doubts, could yet say (Job 27:5-6): “Till I die I will not remove mine integrity from me, etc.” For with this disposition he proved that he did not found his morality on faith, but his faith on morality: in such a case, however weak this faith might be, yet it alone is of a pure and true kind, i.e. the kind of faith that founds not a religion of supplication, but a religion of good life conduct. (MpVT, AA8:267; translation amended)

Job remains in the “midst” of his doubts just as before, if anything with the counterpurposiveness in creation standing out that much more starkly, yet we are told he retains a “pure and true” faith in the face of such doubts. This is not an outright fideism, i.e. faith at the expense of reason, as Kant does offer a clear rational justification for an agent’s faith in a moral God: the only faith “of a pure and true kind,” he says, is “faith [founded] on morality” (MpVT, AA 8:267). *Moral willing itself* is supposed to justify faith as Kant wants to have it, which is no fideism but faith built upon practical reason. Even so, the problem remains that authentic theodicy doesn’t seem to *accomplish* much, that it is all vexation and no resolution. Both the plaintiff against the highest wisdom as well as the self-appointed divine defender, to go back to the trial that begins the essay, might both be read to also live moral lives in their own way, and their interests read as legitimate: indeed, the need

⁴⁰ Gressis 2018 71-3 uses these passages to directly link the “Failure” to the *Religion*, specifically in the notion that it is only from the point of view of moral faith that we can properly “reason” about evil at all.

⁴¹ Kant repeats this in the *Metaphysics of Morals*: “Moral cognition of oneself, which seeks to penetrate into the depths (the abyss) of one’s heart which are quite difficult to fathom, is the beginning of all human wisdom” (MS, AA 6:441).

to sincerely call evil by its name, as well as the need to uphold moral wisdom,⁴² are both aspects of the same moral faith. And if the Kantian response to the “problem of evil” is simply *to live and act morally*, still it seems a hard thing to carry out a moral life in a partly counterpurposive world with all the same doubts present and unanswered. Job may have been “unwise” to start, but if he was already living a moral life to begin with then what has “wisdom” actually benefitted him? All of these parties seem owed better.

If it cannot be enough, as a resolution of the rational interests at issue, to simply show the moral agent the “inscrutability” of the wisdom of creation and then just insist that they keep on in their moral duty, then negative wisdom must have a further function. If faith cannot properly be “founded” on anything beyond morality, it remains that it can become *reflective*.⁴³ Job and his friends, the interests of the two sides of theodicy (and the debate they give rise to), demand some sort of rational resolution. The remaining hope for satisfaction, once we rule out approaches like the doctrinal theodicies, lies in the *self-critique* of those interests themselves regarding exactly what they consist in—or, more broadly, through faith’s becoming self-aware. But as we have seen, faith’s self-consciousness is exactly what *sincerity* is (and it is exactly why Kant emphasizes it with Job). Authentic theodicy, rather than eliminating doubts, is supposed to “handle” the agent’s doubts by transforming their *relation* to their doubts in the context of “a religion of good life conduct,” doing so by progressing already-sincere moral faith into the “conviction of one’s un-knowing,” i.e. negative wisdom. In this sense, negative wisdom—“insight into the necessary limitation of what we may presume with respect to that which is too high for us” (MpVT, AA 8:263)—should be the *culmination* of sincerity, i.e. of faithful self-consciousness.

Thus at the same time as it represents a lived apprehension of divine unknowability, negative wisdom must also be a kind of *self-appropriation*, and in that way satisfy the interests of reason. To the same extent that Job and other sincere moral agents grasp that the highest wisdom is “impenetrable” for our reason, so they must also grasp themselves within that same context—precisely as *moral agents*. Authentic theodicy, the “handling of all objections against divine wisdom” (MpVT, AA 8:264, translation amended), is initially defined as an “*interpretation* of nature insofar as God announces his will through it” and one “made by the law-giver himself,” whether that be God or practical reason. Thus the moral law, and the nature of the free human being who honors it, is at the center of the issue; such theodicy reveals the wisdom of creation, but that in turn points back to *us* and our own freedom. If an account for the presence of the counterpurposive will never be forthcoming, still negative wisdom may provide a *frame* for the moral agent to understand themselves and their own doubts.⁴⁴ The conviction of Job’s un-knowing is at the same time a conviction of his own faith, freedom, and morality—that he will, so to speak, never find more certainty than when he wills his moral projects. Thus the moral agent, in grasping the inscrutability of creation and their own freedom within it, discovers that their doubts and misgivings concerning the wisdom of the world are not so much problems that need solving as confirmations of their faith and the righteousness of their moral enterprise. Even if the moral agent lacks insight into how creation and moral wisdom can ever be reconciled, they can still *rely upon* that unity in their practice: if the evils of the world become excessively troubling then it is the evil state of the world that requires fixing, and not the faith of the agent nor their theology.⁴⁵

So it is in this sense that negative wisdom—as the supplement to, and completion of, a moral faith that is already sincere—can indeed provide satisfaction for reason, at least in the form of self-aware good life conduct. Kant in no way intends to leave the interests of reason hanging, so far as the parties at issue are self-aware enough to appropriate those interests within the context of the broader project of practical reason. “Unwise” plaintiffs may consider the presence of the counterpurposive as good reason to reconsider the righteousness of divine wisdom, just as “unwise” defenders may see unchecked and unanswered doubts as

⁴² Here I read such divine defenders more charitably than Kant does. Regarding Job’s friends, Kravitz suggests: “[T]hey are not driven by an evil motive, but maybe they are over motivated” (Kravitz 2021, 362).

⁴³ Andrew Chignell’s contribution to this volume thematizes the contrast between reflective and unreflective rational belief or faith. Kant himself explicitly uses the term “reflective faith,” and not just rational faith, at the end of the first part of the *Religion*, and it seems to me that this concept—of a faith that is *conscious* of itself—points back to the conclusions of this earlier work: “Reason, conscious of its impotence to satisfy its moral needs, extends itself to extravagant ideas which might make up for this lack, though it is not suited to this enlarged domain... And if in the inscrutable field of the supernatural there is something more than [reason] can bring to its understanding, which may however be necessary to make up for its moral impotence, reason even counts on this something being made available to its good will even if uncognized, with a faith which (with respect to the possibility of this something) we might call *reflective*, since the *dogmatic* faith which announces itself to be a *knowledge* appears to reason dishonest or impudent...” (RGV, AA 6:52).

⁴⁴ On the “Failure” as a case of practical self-reflection, see Lorini 2017 122-127. Additionally, it seems to me that here the concerns of the “Failure” essay in fact dovetail with those of the “righteous Spinoza” passage of KU, AA 5:452-3, albeit beginning from very different directions. Once again, I broadly agree with Esser 2016 that the problem the “limited” moral atheist faces in that passage is not *directly* one of psychological despair and loss of motivation—although that can certainly be a result—but that they cannot make sense of themselves as who they are, as a free being bound to the law, within the context of a purposeless world. In that case the adoption of moral faith removes the “limitation” by showing *both* that the world can be read as morally purposive *and* that their “inner purposive determination by the moral law” (KU, AA 5:452) is not delusion or insanity but constitutive of their being. And Kant pulls exactly this same trick—the discovery of the wisdom in creation as *at the same time* the appropriation of one’s own being—in the “Failure,” to advance Job from sincerity to negative wisdom. In both texts, in very different ways, self-cognition turns out to be pivotal. To be sure, psychologically-driven readers may disagree with my readings of one or both texts, but the consequence is that the case of Job becomes particularly troublesome. For if authentic theodicy does not remove Job’s doubts against faith—if it *cannot*, on pain of insincerity—then it must become unclear whence Job is supposed to source his motivation, or even what role his faith is supposed to have at all.

⁴⁵ Claudia Welz writes, “No matter if we charge, defend or refuse to defend *God’s* love, justice and power—what is, after all, in question is *our* faith, love, justice and powerful or powerless sincerity, *our* solidarity with the suffering, *our* existential commitment. The best theory is that which keeps the wound and the questions open and provokes the continued search for what can heal what is hurt. We need a refusal rather than a refutation of evil” (Welz 2007, 184). Although Welz is somewhat disparaging towards Kant’s account of authentic theodicy, I think this sentiment in fact neatly expresses much of what Kant is suggesting.

immediate dangers to moral faith. However sincere they are in their comportments, neither will get anywhere through speculative argumentation. The right “handling” for the defender of wisdom is the awareness that the best justification of faith is to live morally, even and especially when that is accompanied by doubts; the right “handling” for the plaintiff is the awareness that they not only must but can continue on with all their integrity even in their doubts, well aware that those doubts should in some sense always stand. It is key, therefore, not to conflate (A) the insistence that such parties must carry on in moral action despite having no resolution of their concerns with (B) the proposal that if they understand themselves, they will find that the only resolution to be found must be won within the course of moral action. The most satisfying response to evil is to do good—and that may not be the resolution that such parties were hoping for (it is “vexing” after all), but it is the one to be had.

9. Conclusions

Kant’s admission that counterpurposiveness within the order of nature cannot and should not be argued away—that Job and the plaintiffs against divine wisdom are right, at least with regards to their courage and sincerity—comes with deep consequences for the coherence of his critical system. What Kant had written as an aside in the third *Critique*, that “[t]he moral proof [for the existence of God] would thus always remain in force even if we found in the world no material for physical teleology at all” (KU, AA 5:478), now has more bite: we can no longer unquestionably read nature as compatible with the final end of morality, and we must now acknowledge aspects of the world as standing in tension with moral faith. This truncates the previous ambitions of Kant’s system: there must now be less reliance on compatibilities between nature and morality, and thereby between theoretical and practical reason. Above all the concept of wisdom, the teleological “*Zusammenstimmung*” of everything towards the final ends of reason—the concept which was supposed to finally unify that reason—is relieved of its univocity. In many texts written after the “Failure” it is now *only* God who has wisdom in the proper sense, as in “The End of All Things” (published in the same *Berlinische Monatsschrift* some two and a half years later):

Wisdom, that is, practical reason using means commensurate to the final end of all things—the highest good—in full accord with the corresponding rules of measure, dwells in God alone; and the only thing which could perhaps be called human wisdom is acting in a way which is not visibly contrary to the idea of that wisdom. (EAD, AA 8:336)

But this merely human, *negative* wisdom does not thereby lose its central role, in spite of being divorced from wisdom of a higher kind (and even from systematicity itself). Negative wisdom begins with sincerity (as “formal conscientiousness”), self-awareness concerning one’s faith and morality. And mere sincerity implies no grand systematic ambitions; if anything, its “honesty in openly admitting one’s doubts” and “repugnance to pretending conviction where one feels none” (MpVT, AA 8:266–7) seems more likely to challenge such ambitions. Yet such honesty, Kant now insists, is the surest sign of moral faith and commitment to the final end of practical reason. Negative wisdom is only possible for the faithful agent who is already sincere in this way: authentic theodicy can lay before one’s eyes the inscrutability of creation, and at the same time show the limits of human reason and the interrelations of faith and the counterpurposive, but all of that first relies upon the “honesty (mere simplicity and straightforwardness of mind)” which is (Kant says) “the least that we can possibly require of a good character” (MpVT, AA 8:270). Insight into the limitations of our cognitions and conviction of our un-knowing supplements and completes this moral self-awareness, so that practical reason can “handle” theodical concerns while compromising neither faith nor honest doubts. Wisdom in this new sense becomes central not because it teleologically “tunes together” nature and morality, theory and practice, but because it shows reason that resolution to the questions that burden it can only be found in living well and doing good.

How satisfactory this is depends upon what we, as readers, are looking for. If we go to this essay looking for a better solution to the traditional theodical “problem of evil”—a way to reconcile apparently evil phenomena with a moral God—then Kant ultimately has none to give, since he understands all such ambitions to be in vain at best. What he presents instead is a way for sincere moral agents to “handle” their concerns in an alternative respect, yet one that is still supposed to be rationally satisfying. Mere “negative wisdom” is far indeed from the promises of a moral *Zusammenstimmung* in nature made a year earlier to the righteous Spinoza, but ultimately it serves the same end: to help good people to continue persisting in their projects. A decade earlier Kant wrote that critique, as a “tribunal,” considers both the “rights of reason” (KrV, A751/B779) regarding where and how far its reach can legitimately extend as well as the interests of reason which give rise to such questions in the first place; the critical solution to a philosophical problem “does not consider the question objectively at all, but instead asks about the foundations of the cognition in which it is grounded” (KrV, A484/B512). In this work and in the *Religion* such self-cognition is repeated and renewed, albeit deployed to resolve problems of the heart rather than problems of metaphysics.

10. References

- Allison, H. (1980), “Kant’s Critique of Spinoza,” in R. Kennington (ed.), *The Philosophy of Baruch Spinoza*, Catholic University of America Press, Washington, pp. 199–277
- Beiser, F. (2006), “Moral Faith and the Highest Good,” in P. Guyer (ed.), *The Cambridge Companion to Kant and Modern Philosophy*, Cambridge University Press, Cambridge, pp. 588–629. <https://doi.org/10.1017/CCOL052182303X.018>

- Brachtendorf, J. (2002), "Kants Theodizee-Aufsatz—Die Bedingungen des Gelingens philosophischer Theodizee," *Kant-Studien* 93(1), pp. 57-83. <https://doi.org/10.1515/kant.93.1.57>
- Callanan, J.J. (2019), "Kant on Misology and the Natural Dialectic," *Philosophers' Imprint* 19(47), pp. 1-22.
- Chignell, A. (2007), "Belief in Kant," *Philosophical Review* 116(3), pp. 323-360. <https://doi.org/10.1215/00318108-2007-001>
- Chignell, A. (2023). "Demoralization and Hope: A Psychological Reading of Kant's Moral Argument," *The Monist* 106(1), pp. 46-60. <https://doi.org/10.1093/monist/onac022>
- Duncan, S. (2012), "Moral Evil, Freedom and the Goodness of God: Why Kant Abandoned Theodicy," *British Journal for the History of Philosophy* 20(5), pp. 973-991. <https://doi.org/10.1080/09608788.2012.718860>
- Engstrom, S. (1997), "Kant's Conception of Practical Wisdom," *Kant-Studien* 88(1), pp. 16-43. <https://doi.org/10.1515/kant.1997.88.1.16>
- Eriugena, J.S. (1987), *Periphyseon (Division of Nature)*, trans. I.P. Sheldon-Williams and J.J. O'Meara, Bellarmin, Montreal/Paris.
- Esser, A.M. (2016), "Applying the Concept of the Good: The Final End and the Highest Good in Kant's Third Critique," in T. Höwing (ed.), *The Highest Good in Kant's Philosophy*, De Gruyter, Berlin, pp. 245-262. <https://doi.org/10.1515/9783110369007-016>
- Ferrarin, A. (2015), *The Powers of Pure Reason*, University of Chicago Press, Chicago. <https://doi.org/10.7208/chicago/9780226243290.001.0001>
- Fonnesu, D.L. (2023). "Job's Loneliness. Theodicy and its Failure in Kant's Thought," *Verifiche* 52(2), pp. 1-19. <https://www.verificheonline.net/articoli/jobs-loneliness-theodicy-and-its-failure-in-kants-thought/>
- Fugate, C. (2014), *The Teleology of Reason*, De Gruyter, Berlin. <https://doi.org/10.1515/9783110306484>
- Green, M.K. (1992), "Kant and Moral Self-Deception," *Kant-Studien* 83(2), pp. 149-169. <https://doi.org/10.1515/kant.1992.83.2.149>
- Gressis, R. (2018). "Kant's Theodicy and its Role in the Development of Radical Evil," *Archiv für Geschichte der Philosophie* 100(1), pp. 46-75. <https://doi.org/10.1515/agph-2018-0003>
- Höwing, T. (Ed.) (2016), *The Highest Good in Kant's Philosophy*, De Gruyter, Berlin.
- Husserl, E. (1994) *Early Writings in the Philosophy of Logic and Mathematics*, trans. D. Willard, Kluwer, Dordrecht. <https://doi.org/10.1007/978-94-015-8147-9>
- Huxford, G. (2020), *Kant and Theodicy: A Search for an Answer to the Problem of Evil*, Lexington Books, Lanham.
- Kant, I. (1900-1980), *Kant's gesammelte Schriften* (Vols. 1-29), ed. Königlich Preussischen Akademie der Wissenschaften, Reimer/De Gruyter, Berlin.
- Kant, I. (1992-), *The Cambridge Edition of the Works of Immanuel Kant*, ed. P. Guyer & A.W. Wood, Cambridge University Press, Cambridge.
- Kleingeld, P. (1998), "Kant on the Unity of Theoretical and Practical Reason," *The Review of Metaphysics* 52(2), pp. 311-339
- Kravitz, A. (2020), "Kant's Conception of Theodicy and his Argument from Metaphysical Evil against it," *Archiv für Geschichte der Philosophie* 102(3), pp. 453-476. <https://doi.org/10.1515/agph-2020-1017>
- Kravitz, A. (2021), "Divine Güte, Divine Güte: Kant on an Ancient Query," *Sophia* 60, pp. 349-363. <https://doi.org/10.1007/s11841-021-00825-9>
- Lema-Hincapié, A. (2006), *Kant y la Biblia: Principios kantianos de exégesis bíblica*, Anthropos, Barcelona.
- Lorini, G. (2017). "Till I Die I Will not Remove Mine Integrity from Me. On Kant's 'Anthropological' Theodicy," in N.S. Madrid & H. Robert (eds.), *Kant's Shorter Writings. Critical Paths outside the Critiques*, Cambridge Scholars Publishing, Newcastle upon Tyne, pp. 116-131.
- Munzel, G.F. (1999), *Kant's Conception of Moral Character*, University of Chicago Press, Chicago.
- Munzel, G.F. (2012), *Kant's Conception of Pedagogy*, Northwestern University Press, Evanston.
- Otten, W. & Allen, M. (Eds.) (2014), *Eriugena and Creation*, Brepols, Turnhout.
- Noller, J. (2021), "The Logic of Illusion: Kant on the Reasons of Error," *Theoria* 87(6), pp. 1468-1480. doi:10.1111/theo.12350
- Noller, J. (2022), "Rationalizing: Kant on Moral Self-Deception," *SATS* 23(2), pp. 175-189. <https://doi.org/10.1515/sats-2020-0011>
- Palmquist, S.R. (1992), "Does Kant Reduce Religion to Morality?," *Kant-Studien* 83(2), pp. 129-148. <https://doi.org/10.1515/kant.1992.83.2.129>
- Papish, L. (2018), *Kant on Evil, Self-Deception, and Moral Reform*. Oxford University Press, New York. <https://doi.org/10.1093/oso/9780190692100.001.0001>
- Pasternack, L. (2015), "Kant's 'Appraisal' of Christianity: Biblical Interpretation and the Pure Rational System of Religion," *Journal of the History of Philosophy* 53(3), pp. 485-506. <https://doi.org/10.1353/hph.2015.0042>
- Pasternack, L. (2023), "Kant's Moral Argument and the Problem of Evil : Authentic Theodicy and the Sincerity of Faith," in I. Goy (ed.), *Kant on Proofs for God's Existence*, De Gruyter, Berlin, pp. 139-160.
- Reath, A. (1988), "Two Conceptions of the Highest Good in Kant," *Journal of the History of Philosophy* 26(4), pp. 593-619. <https://doi.org/10.1515/10.1353/hph.1988.0098>
- Schulte, C. (1991), "Zweckwidriges in der Erfahrung. Zur Genese des Mißlingens aller philosophischen Versuche in der Theodizee bei Kant," *Kant-Studien* 82 (4), pp. 371-396. <https://doi.org/10.1515/kant.1991.82.4.371>
- Stevenson, L. (2003), "Opinion, Belief or Faith, and Knowledge," *Kantian Review* 7, pp. 72-101. <https://doi.org/10.1017/S1369415400001746>
- Sticker, M. (2017), "Kant's Criticism of Common Moral Rational Cognition," *European Journal of Philosophy* 25(1), pp. 85-108. <https://doi.org/10.1111/ejop.12197>

- Sussman, D. (2010), "Something to Love: Kant and the Faith of Reason," in B.J.B. Lipscomb & J. Krueger (eds.), *Kant's Moral Metaphysics*, De Gruyter, Berlin, pp. 133–49. <https://doi.org/10.1515/9783110220049.3.133>
- Yovel, Y. (1980), *Kant and the Philosophy of History*, Princeton University Press, Princeton.
- Welz, C. (2007), "Reasons for Having No Reason to Defend God: Kant, Kierkegaard, Levinas and their Alternatives to Theodicy," in H.M. Vroom (ed.), *Wrestling with God and with Evil*, Rodopi, Amsterdam, pp. 167-186.
- Welsch, M. (2019), "Kant über den Selbstbetrug des Bösen," *Kant-Studien* 110(1), pp. 49-73. <https://doi.org/10.1515/kant-2019-0002>
- Wood, A. (1970), *Kant's Moral Religion*. Cornell University Press, Ithaca.