This is a pre-publication draft. To cite this paper, please use the published version:

<https://www.doi.org/10.1515/agph-2020-0003>

**Bolstering the Keystone:**

**Kant on the Incomprehensibility of Freedom**

Tim Aylsworth

“Now, the concept of freedom, insofar as its reality is proved by an apodictic law of practical reason constitutes the *keystone* of the whole structure of a system of pure reason, even of speculative reason . . .” (Kant, *Critique of Practical Reason*, 5:3–4).

“[F]reedom insofar as it is the basis of morality *cannot be comprehended*, and insofar as it can be comprehended it cannot serve as the foundation of morality” (*Kraus’s* *Review of Ulrich’s* Eleutheriology, 8:458).[[1]](#footnote-1)

*1. Introduction*

Kant’s concept of transcendental freedom has puzzled generations of commentators and given rise to some intractable controversies. Indeed, the exegetical and philosophical difficulties are so severe that his view has been described as “incoherent,”[[2]](#footnote-2) “metaphysically preposterous,”[[3]](#footnote-3) and “worthless”[[4]](#footnote-4) — perhaps a valiant effort but ultimately “a hopeless failure.”[[5]](#footnote-5) To make matters worse, there is no consensus on what the view even is. Some readers, lending Kant a charitable hand, have jettisoned much of the theory’s apparent baggage and suggested that the view might be nothing more than the ontologically unambitious claim that when we engage in moral deliberation we are justified in adopting a standpoint from which we think of ourselves as free.[[6]](#footnote-6) Others have defended a metaphysical interpretation, claiming that timeless noumenal choices somehow determine states of affairs in the phenomenal world. Proponents of this view maintain that every rational being has an intelligible character, ontologically distinct from its empirical counterpart, and they argue that this explains how an action in the empirical world could be a consequence of both an uncaused, spontaneous choice and deterministic laws of nature.[[7]](#footnote-7)

Despite the many disputes, the one thing that everyone seems to agree on is that whatever Kant is trying to say about freedom, it is fraught with problems. And this should be cause for concern given the singular importance that he attaches to this concept. Freedom has the privilege of being the “keystone” in the arch of his critical philosophy. It occupies a unique conceptual space at the intersection of practical and theoretical reason,[[8]](#footnote-8) somehow holding the entire edifice together.[[9]](#footnote-9) Morality, as Kant understands it, requires freedom as a necessary condition: it is the “*ratio essendi* of the moral law” (*KpV* 5:5). The special status assigned to autonomy is what makes the good will “shine like a jewel” — a sapphire in the mud of heteronomy (*G* 4:394).

Unfortunately, Kant’s doctrine of transcendental freedom, which underpins the concepts of autonomy and freedom in his practical philosophy, is notoriously problematic.[[10]](#footnote-10) So when we consider the array of problems surrounding Kant’s theory of freedom and the many philosophical puzzles that beset scholars who try to make sense of the view, it should come as a surprise that relatively little attention has been given to his many warnings that freedom is incomprehensible.[[11]](#footnote-11) He makes this point in several places, but it is stated clearly and definitively in the *Groundwork*:

**(1)** Philosophy must therefore assume that no true contradiction will be found between freedom and natural necessity in the very same human actions, for it cannot give up the concept of nature any more than that of freedom. Nevertheless, this seeming contradiction must be removed in a convincing way, *even though we shall never be able to comprehend how freedom is possible* (4:456; emphasis added).

What exactly are we to make of this claim? Kant spills a great deal of ink explaining the details of transcendental freedom, arguing for its compatibility with determinism, and highlighting its crucial significance for his practical philosophy. Why then would he confess that we will never be able to comprehend how it is possible?

My aim in this paper is to answer that question by presenting a novel explanation and defense of Kant’s take on the incomprehensibility of freedom, thereby filling a lacuna in the literature.[[12]](#footnote-12) I conclude that when it comes to theoretical speculation about freedom, the following things are beyond our comprehension: (1) the exact nature of the relation between intelligible and empirical character and (2) determinate, substantive facts about the noumenal self and its free agency as a spontaneous first cause. The strictures of the first *Critique* proscribe the possibility of theoretical cognition of things in themselves. And although the fact of reason establishes the reality (and *a fortiori* the real possibility) of freedom, Kant insists that this is “only for practical use” (*KpV* 5:6). On my view, our practical cognition of freedom, which springs from our awareness that we are bound by the moral law, tells us *that* we are free, but it does not justify any speculation about *how* that freedom works.

It will become clear in what follows that my conclusion is rather contentious. If Kant believes that the details of transcendental freedom are beyond our comprehension, then there are some arguments in the scholarly literature that he might regard as misguided — precisely the kind of dialectical transgressions that he warns us to avoid. Perhaps the most notable case comes up in the debate concerning moral motivation and the role of moral feeling. If I am right about his pessimism regarding our understanding of freedom, then it is Kant’s position that there are certain things we can never know about how an agent acts out of respect for the moral law. Similarly, freedom is occasionally invoked as a bargaining chip in the interpretive debate over transcendental idealism, but I believe the incomprehensibility of freedom undermines the legitimacy of some of these arguments.

The implications of my conclusion, however, are not entirely negative. On the positive side, if Kant is right that there are principled reasons why freedom is beyond our comprehension, then this would lift an undesirable explanatory burden off the shoulders of Kant’s critical philosophy. It would be a boon for Kantians if they could maintain their lofty claims about the unique, elevated status of rational agency without being required to explain all the details of transcendental freedom. It would be especially nice if the reasons for the incomprehensibility thesis are indeed principled, as I believe they are, and it is not just special pleading: a case of Kant’s practical interests compelling us to ask theoretical reason to make an ad hoc exception for the crown jewel of his moral philosophy.

The structure of my paper is as follows. In the next section, I explain Kant’s “incomprehensibility thesis” — the claim that the conditions of cognition give rise to epistemic limitations which make it impossible for us to grasp all the details of transcendental freedom. In section 3, I outline some of the negative implications of this conclusion, highlighting questions and debates which, if I am right about the incomprehensibility of freedom, will turn out to have no definitive answer. I then turn to the positive implications, explaining how freedom creates value *sui generis* and why we are better off acknowledging its incomprehensibility. Although it might seem disappointing at first, I think that this facet of reason’s “peculiar fate” should ultimately be seen in a positive light. I conclude by briefly considering some objections and offering replies.

*2. The Incomprehensibility of Freedom*

My aim in this section is threefold. First, I present textual evidence showing that Kant is committed to the idea that certain things about transcendental freedom are incomprehensible.[[13]](#footnote-13) My second aim is to explain exactly what it is about freedom that Kant thinks is beyond our comprehension. He surely does not mean that we cannot comprehend *anything* about freedom. We must be able to understand, for instance, how it is logically consistent with natural necessity, as he argues in the Third Antinomy.[[14]](#footnote-14) I will therefore try to distinguish between the things that Kant thinks we can and cannot know about freedom. My third goal is to argue that he has principled reasons for the incomprehensibility thesis. It might be tempting to think that the incomprehensibility of freedom is a last resort, a subterfuge Kant employs after failing to give a satisfactory theoretical explanation for something that his practical philosophy requires us to believe. I think we should resist that temptation by coming to see how the incomprehensibility thesis follows directly from Kant’s epistemological commitments. I have divided this section into two parts. First, I show what Kant said about the incomprehensibility of freedom, highlighting the most salient points. I identify what I take to be the four principal themes of Kant’s comments on the incomprehensibility of freedom. These pertain to both the content of and the reasoning behind the Kant’s claim. I elaborate on these points in section 2.2 where I synthesize them into an argument for the incomprehensibility thesis.

*2.1* *Kant on the Incomprehensibility Thesis*

There can be little doubt about Kant’s commitment to the incomprehensibility of freedom; the textual evidence is overwhelming. In his published works alone, it shows up more than twenty times.[[15]](#footnote-15) Here are two of the most perspicuous instances from the *Groundwork*:

**(2)** But reason would overstep all its bounds if it took it upon itself to *explain how* pure reason can be practical which would be exactly the same task as to explain *how freedom is possible*.

For we can explain nothing but what we can reduce to laws the object of which can be given in some possible experience. Freedom, however, is a mere idea, the objective reality of which can in no way be presented in any possible experience; and because no example of anything analogous can ever be put under it, it can never be comprehended or even only seen. (*G* 4:458–59)

**(3)** But it is quite beyond the capacity of any human reason to explain *how* pure reason, without any other incentives that might be taken from elsewhere, can be of itself practical, that is, how the mere *principle of the universal validity of all its maxims as laws* (which would admittedly be the pure form of a pure practical reason), without any matter (object) of the will in which one could take some interest in advance, can of itself furnish an incentive and produce an interest that would be called purely *moral*;it is impossible for us to explain in other words, *how pure reason could be practical*, and all the pains and labors of seeking an explanation of it are lost.

It is just the same as if I tried to fathom how freedom itself as the causality of a will is possible. For then I leave the philosophic ground of explanation behind and I have no other. I might indeed revel in the intelligible world, the world of intelligences, which is still left to me; but even though I have an *idea* of it, which has its good grounds, yet I have not the least *cognizance* of it nor can I ever attain this by all the efforts of my natural faculty of reason. (*G* 4:461–62)

These passages from *Groundwork* IIIcontain several illuminating clues about both the meaning of and the reasoning behind the incomprehensibility thesis. First, it is clear that the conditions of cognition are part of the story. The intelligible self, which possesses transcendental freedom, resides in the noumenal realm, placing it beyond the limits of possible experience and cognition. And as Kant says here, “[W]e can explain nothing but what we can reduce to laws the object of which can be given in some possible experience.”[[16]](#footnote-16) This partially explains the incomprehensibility. We have an idea(*Idee*) of the intelligible self as a free cause, but we have no cognizance (*Kenntnis*) of it. This tracks a distinction that Kant makes in the *Critique of Pure Reason*: the difference between our ability to think (*denken*) of something and to cognize it (*erkennen*).[[17]](#footnote-17) When it comes to thinking about freedom of the intelligible self, no object is represented through sensible intuition because nothing is given to us except through the forms of sensibility — space and time. The intelligible self, however, is not in space and time. Being outside of space and time is the very thing that exempts it from the universal causal necessity of the empirical world. Thus, when we think about transcendental freedom, we are not dealing with a singular, immediate representation of an object given through intuition. It is a mere idea which “signifies only a ‘something’ that is left over when I have excluded from the determining grounds of my will everything belonging to the world of sense” (*G* 4:462).[[18]](#footnote-18) We can revel in thoughts of the intelligible world, but when we do, we fall short of comprehension as we are no longer on the terra firma of cognition.[[19]](#footnote-19)

Furthermore, transcendental freedom requires “an **absolute** causal **spontaneity**” — a causality “through which something happens without its cause being further determined by another previous cause” (A446/ B474). By contrast, our cognition of objects of appearance necessarily involves a causal nexus; nothing happens unless it was determined by a previous state “in accordance with the law of the connection of cause and effect” (B232). We cannot represent transcendental freedom by means of a singular, immediate representation because the conditions of cognition are too restrictive. Cognition requires that an object be given, and no object is given in this case.[[20]](#footnote-20) Thoughts are not limited in this way. Kant says, “I can think whatever I like, as long as I do not contradict myself i.e., as long as my concept is a possible thought, even if I cannot give any assurance whether or not there is a corresponding object somewhere within the sum total of all possibilities” (Bxxvi). In the case of freedom, we merely apply the pure (i.e., unschematized) concept of causality to the idea intelligible self, thinking of it as the ground of certain appearances in the empirical world. But this does not constitute a theoretical cognition. And that places a limitation on our knowledge and comprehension of freedom. This is the first recurrent theme to note about the incomprehensibility thesis. I will refer to it as the **cognition constraint**: our ability to comprehend freedom is constrained by the conditions of cognition.[[21]](#footnote-21)

Second, in the above excerpts from the *Groundwork*, Kant says that explaining “how pure reason can be practical” is “just the same as” explaining how freedom is possible; it is “exactly the same task.” This gives us some additional information about the content of the incomprehensibility thesis. To see why freedom is incomprehensible, we need to understand what Kant means when he says that we cannot explain how pure reason can be practical. In short, the issue is that we cannot comprehend how morality could interest us at all: “The subjective impossibility of *explaining* the freedom of the will is the same as the impossibility of discovering and making comprehensible an *interest* which the human being can take in moral laws . . .” (*G* 4:460). To put it in Kant’s terms, we cannot understand how the pure form of universal law, without the help of empirical incentives (*Triebfedern*), can produce an interest (*Interesse*) in us and serve as the determining ground of the will.[[22]](#footnote-22) Explaining this requires a brief discussion of Kant’s theory of agency.

Ordinarily when we take something as a reason for action, we have an interest in the pursuit of some end (e.g., money, good health, approval from others, etc.); in these instances, we are driven by an empirical incentive. These ends are at the heart of most hypothetical imperatives. For example: “He must work and save in his youth in order to not want in his old age” (*KpV* 5:20). In such cases, the “matter of the practical principle” (i.e., the desired end) is the determining ground of the will. Thus, the prescription has a normative force only insofar as the end is desired. The young man could dispense with this imperative if he “does not hope to live to old age, or thinks that in case of future need he can make do with little” (*KpV* 5:20). It is precisely because of this contingency that Kant argues that practical principles grounded in the desire of some object, “can furnish no practical laws” (*KpV* 5:21). Rather than acting in the pursuit of some end, acting out of respect for the moral law requires us to make the “form of law” the determining ground of the will. But we are unable to comprehend how an idea of pure reason, something that is entirely severed from the empirical world, could generate an interest in us. How can we take an interest in a judgment of pure reason — something devoid of empirical incentives?

Kant has additional reasons for finding this troubling because he believes it is impossible for us to comprehend how an idea of reason could cause pleasure or displeasure. We cannot understand how a pure faculty, could have an effect on sensibility. Moral feeling is shrouded in mystery. He says that it is “quite impossible to see, that is, to make comprehensible a priori, how a mere thought which contains nothing sensible produces a feeling of pleasure or displeasure; for that is a special kind of causality about which, as about any causality, we can determine nothing whatever a priori but must consult experience alone” (*G* 4:460). Our empirical concept of causality can connect only sensible objects with one another. Such knowledge must be gained through experience. This problem case (pure reason’s effect on sensibility) involves a causal connection between a pure idea, which does not originate in the sensible world, and a sensible feeling of pleasure or displeasure. There is no experience that could give us insight into how this works.

The intelligible world contains no trace of incentives from the world of sense. When we try to think of pure reason “with reference to a pure world of understanding as a possible efficient cause” we inevitably discover that “an incentive must be quite lacking” (*G* 4:462). Kant suggests that we might then suppose that the idea of the intelligible world *itself* “would have to be the incentive or that in which reason originally takes an interest; *but to make this comprehensible is precisely the problem that we cannot solve*” (*G* 4:462; emphasis added). At the outset, the task was to understand how pure reason could be practical; that is, we wanted to understand how a mere idea (*viz*., the form of universal validity of maxims as laws), which comes from a pure realm devoid of empirical incentives, could generate an interest in us and serve as the determining ground of the will. Kant then ends up thinking that the idea of the intelligible world might itself be the incentive. But this is just another pure idea, and the problem comes full circle. This is the second point I want to draw from Kant’s remarks in the *Groundwork*.[[23]](#footnote-23) I call it the **problem of** **pure reason’s practicality**. It comprises several issues. We cannot understand how a human being could take an interest in morality, nor can we comprehend how an idea of pure reason could create feelings of pleasure and displeasure or serve as the determining ground of the will. At the center of this problem is the relation between the intelligible and empirical self. If we could fully comprehend that relation, then we would understand how pure reason could be practical.

Kant underscores these points yet again in the *Critique of Practical Reason*. He emphasizes the connection between the limits on cognition and the incomprehensibility of freedom, and he continues to maintain the material equivalence of our inability to understand freedom and the problem of pure reason’s practicality. He does, however, present new arguments in the second *Critique*, and some of these have interesting implications for the incomprehensibility thesis. It is here that Kant presents his infamous “Fact of Reason” argument, claiming that our immediate awareness of the moral law as binding on our will proves the reality of transcendental freedom.[[24]](#footnote-24) This raises interesting architectonic questions about the disparate interests of practical and theoretical reason. For my purposes, the most pressing of these questions is whether or not practical reason’s extension of our knowledge licenses any theoretical speculation about how freedom is possible. Kant insists that it does not:

**(5)** But because no intuition, which can only be sensible, can be put under this application, *causa noumenon* with respect to the theoretical use of reason is, though a possible, thinkable concept, nevertheless an empty one. But I do not now claim *to know theoretically* by this concept the constitution of a being *insofar as* it has a *pure* will; it is enough for me to thereby only designate it as such a being and hence only to connect the concept of causality with that of freedom (and with what is inseparable from it, the moral law as its determining ground); and I am certainly authorized to do so by virtue of the pure, not empirical origin of the concept of cause, inasmuch as I consider myself authorized to make no other use of it than with regard to the moral law which determines its reality, that is, only a practical use. (*KpV* 5:55–56).

**(6)** But is our cognition really extended in this way by pure practical reason, and is what was *transcendent* for speculative reason *immanent* in practical reason? Certainly, but only *for practical purposes.* For we thereby cognize neither the nature of our souls, nor the intelligible world, nor the supreme being as to what they are in themselves [. . .] But how freedom is even possible and how this kind of causality has to be represented theoretically and positively is not thereby seen; that there is such a causality is only postulated by the moral law and for the sake of it. (*KpV* 5:133)

**(7)** But this extension of theoretical reason is no extension of speculation, that is, no positive use can now be made of it for *theoretical purposes.* For, since nothing further is accomplished in this by practical reason than that those concepts are real and really have their (possible) objects, but nothing is thereby given us by way of intuition of them (which can also not be demanded), no synthetic proposition is possible by this reality granted them. Hence this disclosure does not help us in the least for speculative purposes, although with respect to the practical use of pure reason it does help us to extend this cognition of ours. (*KpV* 5:134)

Kant thinks that practical reason extends our knowledge by showing that we do indeed possess transcendental freedom, but it does not give us any further information about how this kind of freedom is possible. Our knowledge of freedom is strictly mediated by the moral law: “the moral law is the *ratio cognoscendi* of freedom” (*KpV* 5:5). We learn from practical reason *that* we are free, but we do not come to understand *how* we are free. This is the third theme of the incomprehensibility thesis. We have practical reasons to believe that we are free, but we do not have theoretical knowledge of how freedom works.[[25]](#footnote-25) I call this the **theoretical knowledge constraint**.

The textual evidence does not end there; it extends far beyond the *Groundwork* and the second *Critique*, showing up in many places before and after the publication of the first *Critique*. Indeed, the incomprehensibility thesis makes appearances in Kant’s notes and lectures as early as the 1760s. It is present in some of his earliest notes and lectures on the subject, and some of these provide important clues about his reasoning, further explaining why freedom is beyond our comprehension:

**(8)** We know well what proceeds from freedom and its presupposition, and it is also necessary for us to presuppose it. But no one can comprehend the origination of a free action, since it is the beginning of all origination. (R 4180)

**(9)** It will still be hard to comprehend by speculative understanding how a derivative being *<ens derivativum>* can perform original acts *<actus originarios*>; but the reason that we cannot comprehend it lies in our understanding, *for we can never conceive the beginning*,but rather only what happens in the series of causes and effects. But the beginning is the boundary of the series, yet freedom makes wholly new divisions for a new beginning; it is on that account difficult to comprehend.But because the possibility of such freedom cannot be comprehended, it does not yet follow from this that, because we cannot comprehend it, there also could not be any freedom. *But freedom is a necessary condition of all our practical actions.* So just as there are other propositions that we do not comprehend but which presuppose a necessary condition, so are we also independent by the concept of transcendental freedom. (*Metaphysik Pölitz*; 28:270)

**(10)** One can have no insight into the possibility of freedom, because one can have no insight into a first beginning, whether the necessity in existence in general or in freedom in the origination of events. For our understanding cognizes existence through experience, but reason has insight into it if it cognizes it *a priori*, i.e., through grounds (that, namely, which is not necessary in accordance with identity, rather which is posited *realiter*); now there are no grounds for that which is first, thus no insight into it is possible through reason. (R 4338)

These two *Reflexionen* and this excerpt from a 1770s metaphysics lecturerefer to something mentioned above in the context of the **cognition constraint**. It was said there that we cannot cognize “absolute causal spontaneity” because a first beginning runs afoul of the understanding’s concept of causality, which is a condition of the possibility of experience. What is said in **(9)** underscores this claim. Theoretical cognition of freedom thus faces two barriers: no object is given through intuition, and our empirical concept of causality cannot process a first beginning. In **(10)**, Kant takes this even further as he says not only that we cannot cognize a first beginning through *experience*, he also claims that we cannot cognize it *a priori* through reason since a first beginning has no grounds through which it could be comprehended. Neither experience nor reason can give us any insight into the nature of a spontaneous, first cause. It is inscrutable. This is the fourth and final component of the incomprehensibility thesis: I will refer to it as the **first beginning problem**.

Kant continues to bring up the incomprehensibility of freedom throughout his later works. It appears several times in the *Critique of the Power of Judgment*,[[26]](#footnote-26) the *Metaphysics of Morals*,[[27]](#footnote-27) and *Religion within the Boundaries of Mere Reason*.[[28]](#footnote-28) These texts address a variety of different questions, but the four themes of the incomprehensibility thesis continue to resonate: First, we have no theoretical cognition of the intelligible self or the transcendental freedom it possesses. Second, we cannot know how pure reason creates an interest in us, for this would involve determinate knowledge of the intelligible character and its relation to the empirical world. Third, although we have practical cognition through the fact of reason, this knowledge is only “for a practical use.” We know we are free because the moral law binds our will, but this practical cognition gives theoretical reason no insight into the mechanism of transcendental freedom. Fourth, we are unable to comprehend a first beginning. A first cause violates the conditions of the possibility of experience, and it cannot be comprehended a priori through reason.

*2.2. The Argument for the Incomprehensibility Thesis*

Having identified these features of the incomprehensibility thesis, we are now able to formulate the argument for it. For the sake of clarity, I have broken it down into a **main argument** and two **auxiliary arguments.**

**The Main Argument**

1. We can comprehend transcendental freedom only if the following conditions are met:

(a) We can comprehend a spontaneous first cause.

(b) We can comprehend the relation between a subject’s intelligible character and empirical character

(2) Conditions (a) and (b) are not met.

Therefore,

(3) We cannot comprehend transcendental freedom.

**Auxiliary Argument (a)**

1. We can comprehend a spontaneous first cause only if we can either cognize it empirically through experience or cognize it a priori through reason
2. We cannot cognize a spontaneous first cause empirically or a priori.

Therefore,

1. We cannot comprehend a spontaneous first cause.

**Auxiliary Argument (b)**

1. We can comprehend the relation between a subject’s intelligible and empirical character only if the following conditions are met:
2. We can comprehend how an idea of pure reason (*viz.*, the universal validity of maxims as laws) could create an interest in us.
3. We can comprehend how pure reason could create feelings of pleasure or displeasure.
4. Conditions (c) and (d) are not met.

Therefore,

1. We cannot comprehend the relation between a subject’s intelligible and empirical character.

The first premise of the **main argument** is definitional. Transcendental freedom is defined negatively, as “independence from everything empirical and so from nature generally” (*KpV* 5:97) and positively as “an absolute causal spontaneity beginning from itself [von selbst] a series of appearances” (A446/ B474). Condition (a) follows clearly enough from the positive definition, but condition (b) requires some explanation.

The key to the resolution of the Third Antinomy — the crucial distinction Kant makes so that freedom and determinism do not conflict with one another — is that transcendental freedom pertains only to an agent’s intelligible character qua noumenon/thing in itself. The laws of nature, which deterministically cause an agent’s actions, pertain only to her empirical character qua phenomenon/appearance. Noticeably absent from this story, however, is an account of exactly how the intelligible character determines the empirical. But we cannot comprehend how transcendental freedom works unless we comprehend this relation between intelligible and empirical character. Consider, for instance, Kant’s discussion of the malicious lie. The empirical investigation into what caused the lie will discover only natural, empirical causes in accordance with deterministic laws. But Kant says that we can disregard all of that and consider the lie to be a consequence of the agent’s intelligible character:

Now even if one believes the action to be determined by these causes, one nonetheless blames the agent [. . .] for one presupposes that it can be entirely set aside how that life was constituted, and that the series of conditions that transpired might not have been, but rather that this deed could be regarded as entirely unconditioned in regard to the previous state, as though with that act the agent had started a series of consequences entirely from himself. This blame is grounded on the law of reason, which regards reason as a cause that, regardless of all the empirical conditions just named, could have and ought to have determined the conduct of the person to be other than it is. And indeed one regards the causality of reason not as a mere concurrence with other causes, but as complete in itself, even if sensuous incentives were not for it but were indeed entirely against it; the action is ascribed to the agent’s intelligible character: now, in the moment when he lies, it is entirely his fault; hence reason, regardless of all empirical conditions of the deed, is fully free, and this deed is to be attributed entirely to its failure to act. (A555/ B583)

Transcendental freedom, the capacity to act as a spontaneous first beginning, belongs to the intelligible self, yet it manifests its effects in the empirical world. The malicious lie was something that happened in space and time, but if we are to regard it as free “in the moment when he lies,” we must think of it as an effect of agent’s intelligible character. Comprehension of this relation is therefore essential to our comprehension of freedom. And if this is true, then the soundness of the **main argument** depends entirely on the soundness of the auxiliary arguments.[[29]](#footnote-29)

The first auxiliary argument concerns our inability to comprehend a spontaneous first cause. Once again, there is a necessary condition that fails to obtain, but this time it consists of a disjunction between two modes of cognition. We have only two ways that we could come to comprehend a spontaneous first cause: cognizing it through experience or a priori through reason. But these two means of cognition will be of no help in the case of a first beginning. As it was explained earlier, there are two reasons why experience does not give us any insight into a spontaneous, noumenal cause. First, no object is given through sensible intuition. Second, the concept of causality is capable only of connecting objects of experience in accordance with a causal law. Even if *per impossibile* such noumenal objects of intuition were given, we would not be able to cognize them through the understanding’s conceptual schemata. It would be an intuition without a concept, which Kant famously describes as “blind” (A51/ B75). Unlike the understanding, however, reason, does not always require that an object be given. Indeed, one of its distinguishing features is that reason entertains ideas for which no corresponding object of cognition is given (e.g., God, the soul, etc.). But this is precisely what ensnares reason into various traps (the antinomies, paralogisms, and the ideal). Absent any concrete grounds (empirical or a priori), reason has no means by which to cognize anything. It spins its own wheels positing objects where none is given in order to satisfy its demand for completeness.[[30]](#footnote-30) Thus, reason and the understanding cannot provide any insight into the nature of a first beginning because there are no grounds through which such a cognition would be possible. The implications of the **cognition constraint** are clearly at work here.

The second auxiliary argument concerns the relation between an agent’s intelligible and empirical character. As I argued earlier, Kant claims in both the *Groundwork* and the second *Critique*, that comprehending how freedom is possible is “just the same as” understanding how pure reason could be practical. And this is explicated in turn by conditions (c) and (d). In order to comprehend how pure reason could be practical, we would need to know how a pure idea (respect for the moral law) could create an interest in an empirical being, or alternatively, how a pure idea could create feelings of pleasure and displeasure. As Kant repeatedly emphasizes, the dual character of the human being is precisely what makes this incomprehensible:

**(15)** If the human being were merely intellectual, then we could have insight into his power of choice through reason; likewise if he were a *brutum*. But not as a sensible and rational being, since his action is subsequently a *phaenomenon*, but antecedently a *noumenon* under practical laws. (R 4788)

The freedom of a purely rational being is relegated entirely to a single realm: the intelligible. It is a rational being exercising its agency according to the dictates of reason. In such cases, there is no question of the relation between an intelligible act of reason and an empirically determined action in space and time. In the case of human agency, things are not so simple, as Kant explains several times in his lectures:

**(16)** How it is possible to be a cause of appearance without itself being appearance cannot be conceived by any human being, nevertheless, we can conceive from this cause that the mechanism of nature does not conflict with freedom, because as appearance a human being is reckoned to the sensible world, and just the same human being is reckoned as intelligence to the intelligible world. (*Metaphysik* *Pölitz*,1790–91,28:583)

**(17)** But we can cognize things not in themselves, what and how they are in themselves, but rather only in appearance, therefore we also know a human being only as he is exhibited to us in the form of sensibility as phenomenon, therefore his actions, but not his determining grounds: *it is thus impossible to prove the real possibility of the absolute self-determination of a human being, or how a human being freely determines himself and nevertheless at the same time is subjected to natural laws*,for that would necessarily require that freedom would be an object of possible intuition. (*Metaphysik Vigilantius* 29:1020)

By relegating freedom and determinism to separate realms, the Third Antinomy demonstrates their compatibility, but it does not provide a concrete account of *how* the intelligible character determines the empirical. Indeed, it must not provide such an account. To do so would overstep the critical boundaries that Kant defends, and he acknowledges as much:

**(18)** Thus in the judgment of free actions, in regard to their causality, we can get only as far as the intelligible cause, but we cannot get **beyond it**; we can know that actions could be free, i.e., that they could be determined independently of sensibility, and in that way that they could be the sensibly unconditioned condition of appearances. But why the intelligible character gives us exactly these appearances and this empirical character under the circumstances before us, to answer this surpasses every faculty of our reason, indeed it surpasses the authority of our reason even to ask it. (A556–57/ B584–85)

Kant says here that we can know that the malicious lie *could* be free, but we cannot know how the liar’s intelligible character gave rise to this empirical character. The mechanism of transcendental freedom is obscure.

Several commentators have made attempts on Kant’s behalf to identify the mechanism of transcendental freedom — to explain the relation between intelligible and empirical character.[[31]](#footnote-31) But among them, Allen Wood shows perhaps the greatest appreciation for the implications of the incomprehensibility thesis with respect to these suggestions. He argues that Kant’s only task in the Third Antinomy is to show that transcendental idealism creates the logical space for a story in which freedom and determinism do not conflict with one another. It does not commit him to anything else. He says that we should think of a positive account of how transcendental freedom works as being akin to a fairy tale that involves mythical creatures but which contains two non-contradictory propositions expressing freedom and determinism respectively:

It would show that *p* and *q* do not contradict each other. It would not matter in the least to its soundness that the narrative as a whole is false, a pure fiction, talking about fairies, witches, noble-minded handsome princes, and a lot of other things found nowhere in reality and believed in only by contemptibly superstitious people [. . .] The fact that noumena (and noumenal selves) are mentioned in the story also makes no difference. We are no more committed to their existence than to that of fairies or witches [. . .] If Kantian ethics is to remain consistent with Kant’s own views about the uncognizability of the intelligible world, then it ought to read Kant’s story about noumenal freedom in exactly the same way as this fairy tale [. . .] No doubt this fiction leaves us — as Kant himself says — with no positive comprehension whatever of how freedom and natural causality in fact coexist in the real world.[[32]](#footnote-32)

The task of the Third Antinomy is to show how transcendental idealism makes it possible for freedom and determinism to coexist. It does not, and indeed must not, commit us to a positive account of how this works. We cannot know how pure reason has effects in the empirical world, and thus we cannot comprehend how transcendental freedom is possible.

*3. The Implications of the Incomprehensibility Thesis*

The incomprehensibility of freedom has implications for a variety of questions and debates pertaining to Kant’s work — too many to exhaustively explore here. I outline a handful of what I take to be some of the most noteworthy implications, highlighting its impact on particular issues within the secondary literature and its significance for Kantian philosophy more broadly. I have called some of these “negative” in the sense that they concern questions or debates where progress is likely to be limited or even impossible because of the incomprehensibility thesis. I have restricted my focus to three such negative implications: the role that moral feeling plays when we act out of respect for the law, the possibility of conversion, and the extent to which freedom can be used as a bargaining chip in the debate over transcendental idealism. In each of these domains, our inability to comprehend freedom has significant consequences and it places limits on how much we can know about these issues. I have dubbed other implications “positive” in cases where recognizing this limitation provides helpful responses to potential problems or objections that could be (and have been) raised against Kant.

*3.1 Negative Implications*

For more than twenty years, there has been a lively debate between “affectivists,” who claim that motivation to act out of respect for the moral law requires an affective component — moral feeling — and “intellectualists,” who argue that a purely intellectual judgment, without the assistance of any feeling, can motivate an agent to act out of respect for the law.[[33]](#footnote-33) For sympathetic readers of Kant’s moral philosophy, this poses something of a dilemma. On the one hand, Kant’s account requires “that the moral law determine the will immediately” and the “determination of the will” must not be accomplished “by means of a feeling, of whatever kind, that has to be presupposed in order for the law to become a sufficient determining ground” (*KpV* 5:71). This seems to imply that feeling must not play a role in moral motivation. But on the other hand, the idea that intellectual judgment alone can motivate action seems to be at odds with both common sense[[34]](#footnote-34) and many of Kant’s comments on the subject. As Kant famously says in one of his ethics lectures:

When I judge by understanding that the action is morally good, I am still very far from doing this action of which I have so judged. But if this judgement moves me to do the action, that is the moral feeling. Nobody can or ever will comprehend how the understanding should have a motivating power; it can admittedly judge, but to give this judgement power so that it becomes a motive able to impel the will to performance of an action — to understand this is the philosophers’ stone. (27:1428)

Regardless of the impact this passage may or may not have on the debate, it is worth noting that it contains yet another mention of the fact that we will never comprehend the mechanism at issue here.[[35]](#footnote-35) Enlightening as it might be to understand how a judgment of pure reason could move us to act morally, the search for such comprehension will never reach its end point.[[36]](#footnote-36) If my argument above is sound and the **problem of pure reason’s practicality** is a central component of the incomprehensibility of freedom, then this should come as no surprise. We cannot comprehend how pure reason could be practical nor can we understand how pure reason could have an effect on pleasure and displeasure.[[37]](#footnote-37) It would therefore be quite troubling if this is the problem affectivists and intellectualists are trying to solve. Perhaps the more charitable interpretation would be to assume that this is not what they are trying do.[[38]](#footnote-38)

But that leaves us with a difficult question about what their aim is and how to avoid the conclusion that they are trying to explain how pure reason is practical. One way of sidestepping the issue would be to claim that the things we cannot comprehend all lie upstream, in the mysterious noumenal realm of intelligible beings, and that the locus of the debate over moral feeling is to be found downstream in the comprehensible, empirical world. A move of this sort has been suggested,[[39]](#footnote-39) but it sometimes runs into a problem. If we restrict our focus to the empirical world we would indeed avoid both components of the incomprehensibility thesis: we would not need to say anything about the free agency of the intelligible self or about its relation to empirical character. The problem, however, is that by shifting the discussion to the empirical world, we make it impossible to talk about human frailty — one of the issues at the heart of the debate. Affectivists have argued that frailty is a problem for the intellectualist view insofar as it demonstrates that an intellectual judgment of moral rightness is insufficient to motivate action. In the *Religion*, Kant characterizes this kind of frailty as one of the three degrees of evil (along with impurity and depravity).[[40]](#footnote-40) And he makes it clear that this evil concerns the noumenal self and its propensity to subordinate the moral law to inclinations of self-love in the case of our fundamental moral disposition (*Gesinnung*).[[41]](#footnote-41) This results in yet another dilemma. The debate between affectivists and intellectualists either involves asking questions about the intelligible self’s free agency and its relation to empirical character (in which case they are doing something that Kant thinks we should avoid), or they can restrict the discussion to empirical psychology, which would make it impossible to discuss moral agency, frailty, and radical evil in the way that Kant does.[[42]](#footnote-42) But if the incomprehensibility thesis is taken seriously, then perhaps we would avoid some of these questions altogether by heeding Kant’s warning that it is impossible to explain “how pure reason could be practical, and all the pains and labors of seeking an explanation of it are lost” (*G* 4:461). This is not to say that the entire debate is an idle dispute. A live question remains about moral motivation and the phenomenology of respect. I suspect that the solution lies in framing the debate neither in terms of empirical psychology nor with reference to incomprehensible noumenal freedom. As an issue of moral phenomenology, the interesting question lies somewhere between these two horns of the dilemma. But even if we adopt such a middle way, I would still contend that the affectivists go awry when invoking frailty as a flaw of the opposing view. If the objection is that we cannot comprehend how a pure intellectual judgment has motivational force (or why it sometimes fails), then it would seem that the incomprehensibility of freedom (particularly the **problem of pure reason’s practicality**) is the very thing at issue.

The second negative implication I address concerns the possibility of conversion. Kant claims in the *Religion* that there is a “radical innate evil in human nature” that consists in the fact that the human being “is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it” (*Rel* 6:32). The question then becomes how reform is possible. How can we convert our moral disposition from evil to good? Kant himself is perplexed by the question:

How it is possible that a naturally evil human being should make himself into a good human being surpasses every concept of ours. For how can an evil tree bear good fruit? But, since by our previous admission a tree which was (in its predisposition) originally good but did bring forth bad fruits, and since the fall from good into evil (if we seriously consider that evil originates from freedom) is no more comprehensible than the ascent from evil back to the good, then the possibility of this last cannot be disputed. For, in spite of that fall, the command that we ought to become better human beings still resounds unabated in our souls; consequently, we must also be capable of it . . . (*Rel* 6:44–45).

The agent freely chooses to subordinate the moral law to self-love — a choice that corrupts her fundamental moral disposition. Kant then wonders how it would be possible for this evil tree to bear good fruit. If her disposition is utterly corrupted by radical innate evil, how could it be the origin of a conversion? Kant claims that, even if it is incomprehensible (and indeed it is),[[43]](#footnote-43) conversion must be possible given that we “ought to become better,” making this one of several places in the *Religion* where he invokes the “ought implies can” principle.

This has not gotten Kant off the hook with commentators, however. In a recent paper, Samuel Loncar claims that Kant’s argument is flawed. In this excerpt from the *Religion*, Kant seems to be claiming that our fall into evil is incomprehensible but actual (*a fortiori* possible), and if our conversion from evil to good is “relevantly parallel,” then it too is incomprehensible but possible.[[44]](#footnote-44) The problem, Loncar argues, is that the fall and the conversion are importantly distinct and thus the parallelism argument fails. They look similar on the surface, insofar as the fall involves having a good disposition but then adopting an evil maxim, and conversion involves having an evil disposition and subsequently reforming. But this similarity breaks down upon closer analysis. Kant says that our predisposition for good consists in nothing other than our capacity to be good: we are aware of the moral law but we have not yet incorporated it into our fundamental maxim.[[45]](#footnote-45) This is indeed quite unlike conversion which involves reforming a disposition that has *already* adopted an evil maxim. Loncar ultimately concludes that Kant’s view of conversion is to be salvaged by recognizing the necessity of “a self-creating subject who preexists subjection to reason and is indeed responsible for subjecting itself to reason.”[[46]](#footnote-46) Kant rejected Fichte’s self-positing *Ich*, but Loncar thinks Kant’s account of radical evil and conversion is pushed toward this very idea even though it lies “beyond the ken of Kant’s philosophy.”[[47]](#footnote-47)

This suggestion is an interesting one, but it is not the only way out of the bind. The parallel Kant draws between conversion and the fall might not concern their formal structure at all. The relevant parallel might simply be that they are both incomprehensible.[[48]](#footnote-48) Indeed, if my position on the incomprehensibility of freedom is right, then I think we can (and should) avoid asking any questions whatsoever about how conversion is possible. Just as the debate over moral feeling is entangled with the **problem of pure reason’s practicality**, the question of conversion cannot get around the **first beginning problem**. Speculation about the possibility of conversion (and about the fall) necessarily involves asking questions about the spontaneous free agency of the intelligible self. We should not expect to find any definitive answers to these questions. We might be inclined to wonder about them. Indeed, Kant sometimes suggests that we cannot resist this speculation; such is the “peculiar fate of reason.” It is “burdened with questions which it cannot dismiss, since they are given to it as problems by the nature of reason itself, but which it also cannot answer, since they transcend every capacity of human reason” (A vii).

The last negative implication that I will explore concerns the interpretive debate over transcendental idealism. The dispute centers around the question of how to interpret the distinction between appearances and things in themselves. The resolution to the third antinomy depends crucially on this distinction, and so this issue has been a significant concern for commentators who are interested in making sense of Kant’s view of freedom. The intelligible self, the self qua noumenon/thing in itself is free. The empirical self, the self qua phenomenon/appearance, is causally determined by natural laws and thus unfree. Unfortunately, as I mentioned in the introduction, there is no consensus on how to take this distinction, and interpretations are widely divergent. On the one hand, there are those who claim that the intelligible and empirical self are identical, that there is only one world — one set of objects.[[49]](#footnote-49) On the other hand, some deny the identity of the noumenal and phenomenal self, asserting the existence of two, ontologically distinct worlds — two sets of objects.[[50]](#footnote-50)

Those familiar with Kant scholarship are likely to be well acquainted with this debate, and it is not my intention to rehash it here or to defend any particular view. My interest is simply to explore the implications that the incomprehensibility of freedom should have on the debate. It is not uncommon for commentators to invoke freedom as a bargaining chip in this debate, arguing that one view or another fails to make sense of transcendental freedom. While I do not believe that all of these arguments run afoul of the incomprehensibility thesis, some of them do. There is one tactic in particular that should give us pause. The two-standpoint view is sometimes seen as having an advantage over the two-world view on the grounds that the latter either (A) requires (but lacks) a substantive view of how the two worlds relate or (B) presents a substantive view of the relation albeit one that is implausible.[[51]](#footnote-51) Consider the following examples:

This two-standpoint interpretation has at least two important advantages over two-world accounts. First, it allows one to make use of Kant’s insights about freedom from a practical point of view without making substantive metaphysical assumptions. Especially for those primarily interested in Kant as a moral philosopher, this advantage is considerable. Second, *it helps one avoid difficult problems about how different worlds can relate* . . .[[52]](#footnote-52)

For example, the two-worlds account faces the pressing problem of how to explain the relationship between the two worlds to which it appeals. It seems that the “noumenal I” must bear some relation to some of the actions that take place in the phenomenal world of time and space. Otherwise, the fact that the “noumenal I” is free and its actions undetermined would have no relevance to the actions for which we normally take people to be responsible. Yet it is unclear how anything in the noumenal world, distinct from the phenomenal world of space and time, could “affect” that world. Thus, the two-worlds view saddles us with a metaphysical apparatus that seems to raise as many problems as it solves.[[53]](#footnote-53)

On the ontological view [of freedom], the question how the two worlds are related is one which, frustratingly, cannot be answered.[[54]](#footnote-54)

If I am right about the incomprehensibility thesis, then it is Kant’s view that we cannot know the details of the relationship between free noumenal agency and actions in the empirical world. As I see it, the textual evidence on that count is abundant. Thus, if Kant’s distinction is indeed one that involves two ontologically discrete worlds, then it is not incumbent on such a view to commit itself to a view of how the two selves relate. Indeed, it must not commit itself to such a view. What should be offered is a logically possible explanation of how there is no contradiction between freedom and determinism: the noumenal self is free and the phenomenal self is determined. Further details about how an agent’s intelligible character gives rise to her empirical character are unknowable, and we should neither ask for them nor hold interpretive views responsible for providing them.[[55]](#footnote-55) This is not to say that it is always inappropriate to bring up freedom in the transcendental idealism debate. What would make an objection problematic is if it unfairly requires its opponents to perform the impossible task of explaining the relation between the intelligible and empirical self.[[56]](#footnote-56)

If commentators were to take the implications of the incomprehensibility thesis seriously, the landscape of these three discussions would have to change. The debate over moral feeling would have to situate itself more clearly in relation to the incomprehensibility of freedom by explaining what it is they are trying to understand and how exactly it differs from the inexplicable practicality of pure reason. Open questions about the possibility of conversion from evil would not constitute grounds for finding fault with Kant, as we would have to recognize that the nature of conversion is (and ought to be) incomprehensible.[[57]](#footnote-57) And participants in the debate over transcendental idealism would have to abandon objections that demand an explanation of how the free agency of the intelligible self manifests itself in the empirical world.

*3.2 Positive Implications*

Although the incomprehensibility of freedom limits how much we can know about our agency, it also opens a space in Kant’s practical philosophy for claims that are above the reproach of theoretical objections. Some of these include Kant’s celebrated moral views that have been admired by generations of scholars in spite of reservations about the connection between these positions and the seemingly untenable the doctrine of transcendental freedom. Here are two familiar claims from Kant’s moral philosophy, which he thinks depend crucially on transcendental freedom:

1. The moral law has a special, elevated status: its dictates are categorical and overriding.
2. Human beings, in virtue of their capacity for autonomous action, have an unconditional, incomparable worth/dignity (*Würde*).

It is not my intention to offer an explanation or defense of these well-known claims. What I want to show is that, on Kant’s view, both of these positions depend on transcendental freedom. This means that objections to Kant’s view of freedom pose problems for his moral philosophy, and if the incomprehensibility thesis can forestall those objections, then it would act as a safeguard for his practical philosophy.

According to Kant, prescriptions that come from the moral law do not stand on equal footing with those that come from inclinations of self-love. On the contrary, the dictates of morality are given an overriding weight, which is unlike all other driving forces; it infringes on self-love and strikes down self-conceit — nullifying their unwarranted claims on our will (*KpV* 5:73). As Kant sees it, there is chasm separating autonomy and heteronomy, and it consists in the fact that they have distinct grounds. Throughout both the *Groundwork* and the second *Critique*, Kant is at pains to demonstrate that the moral law is of pure (i.e., non-empirical) origin. If morality were tainted with empirical grounds, he thinks it would yield only contingent, subjective prescriptions. For Kant, there are two realms of laws, nature and freedom, and the laws of morality must not come from nature. This is why Kant says that freedom is the *ratio essendi* of the moral law (*KpV* 5:5). Indeed, in a rather remarkable passage from the *Critique of Practical Reason*, Kant says that both the unique respect we have for the moral law and the special status attributed to human beings (in virtue of their capacity for morality) are grounded in freedom. He asks what origin could be worthy of the name ‘duty,’ which brings forth “a law before which all inclinations are dumb,” and which constitutes the “indispensable condition of that worth which human beings alone can give themselves” (*KpV* 5:86). The answer is transcendental freedom:

It can be nothing less than what elevates a human being above himself (as a part of the sensible world), what connects him with an order of things that only the understanding can think and that at the same time has under it the whole sensible world and with it the empirically determinable existence of human beings in time and the whole of all ends (which is alone suitable to such unconditional practical laws as the moral). It is nothing other than *personality*, that is, freedom and independence from the mechanism of the whole of nature, regarded nevertheless as also a capacity of a being subject to special laws . . . (*KpV* 5:86–87).

Although it might already be evident from what he says here, Kant goes on to clarify that transcendental freedom alone can justify these lofty claims of his practical philosophy.[[58]](#footnote-58) Other views of freedom will not suffice. He offers a particularly derisive criticism of Leibnizian compatibilism, which labels some causally necessitated actions “free” on the grounds that their determining grounds were “within the acting being.” Kant mocks this view of freedom saying that it is “nothing better than the freedom of a turnspit, which, when once it is wound up, also accomplishes its movements of itself” (5:97).

This adds considerable weight to the criticisms mentioned in the introduction.[[59]](#footnote-59) If Kant’s view of freedom is indeed an incoherent, hopeless failure, then so is a significant portion of his moral philosophy. Without the pure ground of transcendental freedom, both the elevated status of human beings and the unique respect for the moral law would be undermined. There would not be anything special about reason after all; it could turn out to be the slave of the passions, just as Hume had argued. An objection of this kind could be quite troublesome for those who are seriously interested in Kant’s ethics.

One popular way of dealing with this problem is to adopt the two-standpoint view and thus avoid any metaphysical claims concerning the relation of free choices and the empirical world. It is easy to see why this might be attractive. If we were faced with a choice between accepting the two-standpoint interpretation or committing to an absurd metaphysical view of freedom, then the two-standpoint approach would surely be the way to go. I have argued, however, that the incomprehensibility of freedom exposes this as a false dilemma. There is a third alternative. Kant’s distinction between things in themselves and appearances might be a metaphysical one, but we do not have to commit ourselves to any particular explanation of how the intelligible character gives rise to its empirical counterpart. Indeed, we should not commit ourselves to such a view. To do so would overstep the boundaries established in the first *Critique*. Kant believes that theoretical reason must be satisfied with mere logical possibility; anything beyond that is unwarranted speculation. This is why Kant is so tentative in his conclusion to the Third Antinomy, claiming that he has not shown the reality or even the real possibility of freedom.[[60]](#footnote-60) When we are confronted with certain objections about the plausibility of transcendental freedom, perhaps the best response would be to concede agnosticism and acknowledge that we do not really have a positive view of how freedom is possible.[[61]](#footnote-61) Kant’s view is that we are immediately conscious of the moral law, and this shows the reality of transcendental freedom. He does not think we need to do anything further to respond to critics of freedom:

**(19)** So it is not worth the trouble, either, to refute all the objections leveled against freedom. In this determination of the consciousness of freedom, namely through the categorical imperative, the main question is always the one already illustrated: How is such a categorical imperative possible? This is the most difficult point, since it can neither be proved nor rendered comprehensible; the possibility rests solely on the presupposition of freedom. (27:507)

If Kant is right and we have principled reasons for dismissing criticisms of the plausibility of transcendental freedom, then the incomprehensibility thesis should be seen not merely as a limitation on our knowledge but as a comforting palliative that assuages some of the deepest worries about Kant’s critical philosophy. It safeguards not only the two ethical claims mentioned above, it underpins the value of the world itself.[[62]](#footnote-62)

*4. Conclusion*

I would like to conclude by briefly considering objections to the views I have suggested in this paper. I begin with objections to my central argument from section 2, and then I address potential reservations about the implications I raised in section 3. The first objection concerns Kant’s claims about practical cognition and the extent to which they might undermine the argument I have given for the incomprehensibility thesis. One central component of Kant’s conclusion that we cannot comprehend transcendental freedom is his claim from the first *Critique* that we have no theoretical cognition of things in themselves. But in the *Critique of Practical Reason* he seems to walk this back a bit as he suggests that our consciousness of the moral law gives us some insight into the noumenal world, as it demonstrates the reality of transcendental freedom by means of a “practical cognition [*praktische Erkenntnis*]” (*KpV* 5:103).

This objection is not too worrisome, and the reply to it can probably be predicted from my remarks in section 2 regarding the **theoretical knowledge constraint**. Kant repeatedly emphasizes that practical cognition demonstrates the reality of freedom without providing any insight into how it works. This is one of the refrains at the end of the second *Critique*: our cognition is indeed extended into the supersensible but “only *for practical purposes*” (*KpV* 5:133). When we think of people in moral terms — making character evaluations and attributing praise or blame — we must put aside the causal necessity of nature and believe that they could have done otherwise. This means that we are justified in asserting the reality of transcendental freedom. But we cannot make the leap from practical reason’s certainty to theoretical reason’s speculation. As I explained above, Kant’s view is that practical reason licenses the assertion of freedom but “how freedom is even possible and how this kind of causality has to be represented theoretically is not thereby seen; that there is such a causality is only postulated by the moral law and for the sake of it” (*KpV* 5:133).[[63]](#footnote-63)

The second objection is somewhat more troubling. I have argued that Kant can legitimately avoid responding to objections to his doctrine of transcendental freedom on the grounds that he has no positive view of how freedom actually works. The details, whatever they might be, are beyond our comprehension. But I also seem to have committed Kant to a particular metaphysical view of transcendental freedom — one that involves timeless, noumenal agents acting as first causes and serving as the grounds of appearances in the phenomenal world. It looks like I have committed Kant to a metaphysically extravagant view but then made it immune to objection by fiat. I could be accused of claiming both that Kant has no view and that he has a preposterous one. This is likely to come across as exactly the kind of maddening subterfuge I mentioned in the introduction.[[64]](#footnote-64)

My response to this objection is that we must walk a fine line in order to secure a middle way between two untenable positions. On the one hand, there is the view that what Kant says about transcendental freedom should be treated as a total fiction.[[65]](#footnote-65) Every detail he provides about transcendental freedom in his discussion of the Third Antinomy is given as part of a logically possible story, which demonstrates that freedom and determinism are not necessarily incompatible, but Kant is not committed to those details and he does not expect us to accept anything he says there as true. We should pass over transcendental freedom in silence. On the other hand, there is the view that Kant is committed to a particular metaphysical account of transcendental freedom, and he tells us all about it when he resolves the Third Antimony. Neither of these views is particularly attractive.

As palatable as the first interpretation might seem to those who would like to purge Kant of undesirable metaphysical views, it is surely incompatible with what Kant says about the extent of his commitment to transcendental freedom, especially in the second *Critique*. In the Preface, Kant tells us that the reality of transcendental freedom is firmly established “taken indeed in that absolute sense in which speculative reason needed it, in its use of the concept of causality, in order to rescue itself from the antinomy in to which it unavoidably falls . . .” (*KpV* 5:3). Using particularly strong epistemic language, Kant says here that our freedom is proven (*bewiesen ist*). So how can Kant’s view be that we should pass over transcendental freedom in silence if he tells us that his practical philosophy establishes the reality of precisely that kind of freedom which he had described in the first *Critique*?[[66]](#footnote-66)

If the problem with the first view is that it involves too little of a commitment to transcendental freedom, the problem with the second is that it burdens us with too much of a commitment. Earlier, I mentioned four different ways that commentators have suggested in order to make sense of what Kant says in his resolution of the third antinomy.[[67]](#footnote-67) It could be that free noumenal choices ground the laws of nature.[[68]](#footnote-68) Perhaps God chooses to create noumenal selves whose choices conform to the phenomenal laws.[[69]](#footnote-69) Or it could be something else entirely. Kant’s view is that we can never know exactly how this works. I think Wood is right to suggest that we should, in a certain sense, treat these stories as fictions. Their importance lies not in giving us an account of freedom that we are supposed to believe but in showing us that it is at least *logically possible* that transcendental freedom and determinism could be compatible. I think Wood goes too far, however, when he claims that we should dismiss the entire idea of transcendental freedom as a fantasy of transcendent metaphysics.

This leads back to the question that was at issue throughout this paper. Exactly how much does Kant think we can know about transcendental freedom? How much of the extravagant metaphysical view does he really expect us to believe? I think the answer is that we must be satisfied with a very thin description of the concept. Kant is committed to the reality of transcendental freedom; so, at the very least we should have some idea how to define that concept. If we look closely at his various definitions we find a number of recurrent themes, especially the two I highlighted above in the context of the **auxiliary arguments**. Transcendental freedom involves a complete “independence from everything empirical and so from nature generally” (*KpV* 5:97) and an “absolute causal spontaneity beginning from itself a series of appearances that runs according to natural laws” (A446 / B474). Thus, it must involve noumenal agency, and the choices of noumenal agents must have effects in the phenomenal world in accordance with deterministic laws of nature. That is all we can know about transcendental freedom. We cannot comprehend our noumenal agency (how it converts itself from evil, etc.), and we cannot know how our noumenal agency manifests itself in the empirical world. It is akin to my beliefs about caffeine and its effect on my brain. I have *some* concept of how coffee makes me feel more alert, but I do not fully comprehend the biochemical mechanisms that are operating behind the scenes.[[70]](#footnote-70) I know that there must be such a mechanism, but I do not know exactly what it is.

The last objection concerns the negative implications I explored in section 3. I argued there that the incomprehensibility of freedom has not been fully appreciated in some of the scholarly discussions of moral feeling and conversion. On my interpretation, Kant holds the view that we can never know exactly how pure reason is practical or how conversion is possible. But if Kant holds the view that these are cases of transcendent metaphysical speculation, beyond the bounds of comprehension, then why does he spend so much time discussing them? Chapter 3 of the *Critique of Practical Reason* is dedicated entirely to the incentives of pure reason, and Part One of the *Religion* is all about the fall of radical evil and the possibility of conversion. It looks like I must accuse Kant of transgressing the very boundaries that he established.

In the case of conversion, I would concede that Kant is going beyond the boundaries of comprehension, but I think he is fully aware of that. He says outright that both the fall from good into evil and the possibility of conversion are incomprehensible (*Rel* 6:45). And when he does speculate about how a conversion might work, he countenances some possibilities but ultimately admits that is beyond our comprehension.[[71]](#footnote-71) I think this is in keeping with his views on the incomprehensibility of freedom and the peculiar fate of reason. Moral feeling is a bit more complicated. Not every component of moral feeling is beyond our comprehension. Kant’s discussion of it in the second *Critique* begins with a statement of the incomprehensibility thesis.[[72]](#footnote-72) He says we cannot know the ground from which pure reason creates an incentive and that his aim in the chapter is merely to explore what happens downstream from this determination: “what it effects in the mind” (*KpV* 5:72). Much of moral feeling is well within our grasp: we can explore the phenomenology of respect, what it feels like when it strikes down self-conceit, the complicated mix of pleasure and displeasure it involves, and so forth.[[73]](#footnote-73) My concern with the debate in the secondary literature is that it vacillates between the comprehensible and incomprehensible components of pure reason’s practicality. I think, just as Kant did, that there is something valuable to be learned from exploring the nature of respect for the moral law, but I believe that discussions of it must be more careful with the boundaries Kant put in place.

The incomprehensibility of freedom could have a variety of such consequences for Kant’s views that I have not explored in this paper. But I do not think the negative implications are the principal upshot of the position I have defended here. I think the most important result is that Kant offers us a viable way of believing in a kind of freedom that gives unparalleled weight and meaning to our actions. The epistemic limitations that he defends in the *Critique of Pure Reason* safeguard transcendental freedom. The incomprehensibility thesis shields us from objections that threaten to undermine our belief in that kind of freedom — skeptical doubts that would throw the “inner value of the world” into question.[[74]](#footnote-74)

References

Allais, Lucy. “Kant’s One World: Interpreting Kant’s ‘Transcendental Idealism.’” *British Journal for the History of Philosophy* 12 (2004): 655–84.

———. *Manifest Reality: Kant’s Idealism and his Realism*. New York: Oxford University Press, 2015.

Allison, Henry. “Empirical and Intelligible Character in the *Critique of Pure Reason*” in *Kant’s Practical Philosophy Reconsidered*, edited by Yirmiyahu Yovel, 1–22. Jerusalem: Springer, 2009.

———. *Kant’s Theory of Freedom*. Cambridge: Cambridge University Press, 1990.

Ameriks, Karl. “Is Practical Justification in Kant Ultimately Dogmatic?” in *Kant on Practical Justification: Interpretive Essays*, edited by Mark Timmons and Sorin Baiasu, 153–75. New York: Oxford University Press, 2013.

———. “Kant’s Deduction of Freedom and Morality”

———. *Kant’s Theory of Mind*. Oxford: Oxford University Press, 2000.

———. “Kant, Hume, and the Problem of Moral Motivation.” *In Kant and the Historical Turn: Philosophy as Critical Interpretation*, 89–107. Oxford: Clarendon Press, 2004.

———. “Kant, Miracles, and *Religion*, Parts One and Two” in *Kant’s* Religion within the Boundaries of Mere Reason: *A Critical Guide*, edited by Gordon Michalson, 137–55. Cambridge: Cambridge University Press, 2014.

Beck, Lewis White. *Commentary on Kant’s* Critique of Practical Reason. Chicago: Chicago University Press, 1960.

Bennett, Jonathan. *Kant’s Dialectic*. Cambridge: Cambridge University Press, 1974.

———. “Kant’s Theory of Freedom” in *Self and Nature in Kant's Philosophy* edited by Allen Wood, 102–112. Ithaca, NY: Cornell University Press, 1984.

Bird, Graham. “The Antinomies: An Entirely Natural Antithetic of Human Reason” in *Kant’s* Critique of Pure Reason: *A Critical Guide*. Edited by James O’Shea. Cambridge: Cambridge University Press, 2017.

———. *The Revolutionary Kant: A Commentary on the* Critique of Pure Reason. Chicago, IL: Open Court Press, 2006.

Bok, Hilary. *Freedom and Responsibility*. Princeton, NJ: Princeton University Press, 1998.

Chignell, Andrew. “Belief in Kant.” *Philosophical Review* 116 (2007): 323–60.

———. “Real Repugnance and Belief in Things in Themselves” in *Kant’s Moral Metaphysics*, edited by Benjamin J. Bruxvoort Lipscomb and James Krueger, 177–210. Göttingen: Walter de Gruyter, 2010.

Engstrom, “The *Triebfeder* of Pure Practical Reason” in *Kant’s* Critique of Practical Reason: *A Critical Guide*, edited by Andrews Reath and Jens Timmerman, 90–118. Cambridge: Cambridge University Press, 2010.

Frierson, Patrick. *Kant’s Empirical Psychology*. Cambridge: Cambridge University Press, 2014.

———. “Two Standpoints and the Problem of Moral Anthropology” in *Kant’s Moral Metaphysics*, edited by Benjamin J. Bruxvoort Lipscomb and James Krueger, 83–110. Göttingen: Walter de Gruyter, 2010.

Grenberg, Jeanine. “In Search of the Phenomenal Face of Freedom” in Kant’s Moral Metaphysics, edited by Benjamin J. Bruxvoort Lipscomb and James Krueger, 111–32. Göttingen: Walter de Gruyter, 2010.

———. “Review: McCarty, *Kant’s Theory of Action*.” *Mind* 119 (2010): 1198–1205.

Grier, Michelle. *Transcendental Illusion*. Cambridge: Cambridge University Press, 2004.

Guyer, Paul. *Freedom, Law, and Happiness*. Cambridge: Cambridge University Press, 2000.

———. *Kant and the Claims of Knowledge*. Cambridge: Cambridge University Press, 1987.

———. *Kant’s System of Nature and Freedom*. Oxford: Oxford University Press, 2005.

———. “Naturalistic and Transcendental Moments” *Inquiry* 50 (2007): 444–64.

———. “Problems with Freedom” in *Kant’s* Groundwork of the Metaphysics of Morals: A Critical Guide, 176–202. Cambridge: Cambridge University Press, 2009.

Henrich, Dieter. “Die Deduktion des Sittengesetzes” in *Denken im Schatten des Nihilismus: Festschrift für Wilhelm Weischedel*, edited by Alexander Schwan, 55–112. Darmstadt: Wissenschaftliche Buchgeselleschaft, 1975.

Herrera, Larry. “Moral Triebfeder.” *Kant-Studien* 91 (2000): 395–410.

Holberg, Erica. “The Importance of Pleasure in the Moral for Kant’s Ethics.” *Southern Journal of Philosophy* 54 (2016): 226–46.

Josifovic, Sasa and Noller, Jörg. *Freiheit nach Kant: Tradition – Rezeption – Transformation – Aktualität*. Brill-Verlag (Forthcoming).

Kain, Patrick. “Practical Cognition, Intuition, and the Fact of Reason” in *Kant’s Moral Metaphysics*, edited by Benjamin J. Bruxvoort Lipscomb and James Krueger, 211–232.

Kant, Immanuel. *Critique of the Power of Judgment*, trans. Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press, 2002.

———. *Critique of Practical Reason* in *Practical Philosophy*,trans. Mary J. Gregor. Cambridge: Cambridge University Press, 2008.

———. *Critique of Pure Reason*, trans. Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, 1998.

———. *Groundwork of the Metaphysics of Morals* in *Practical Philosophy*, trans. Mary J. Gregor. Cambridge, U.K.: Cambridge University Press, 2008.

———. *Lectures on Ethics*, trans. Peter Heath. Edited by Peter Heath and J.B. Schneewind. Cambridge: Cambridge University Press, 1997.

———. *Lectures on Logic*, trans. J. Michael Young. Cambridge: Cambridge University Press, 1992.

———. *Lectures on Metaphysics*, trans. Karl Ameriks and Steve Naragon. Cambridge: Cambridge University Press, 1997.

———. *Metaphysics of Morals* in *Practical Philosophy*, trans. Mary J. Gregor. Cambridge, U.K.: Cambridge University Press, 2008.

———. *Notes and Fragments*, trans. Curtis Bowman, Paul Guyer, and Frederick Rauscher. Cambridge: Cambridge University Press, 2005.

———. *Religion within the bounds of Mere Reason* in *Religion and Rational Theology*, trans. Allen Wood and George di Giovanni. Cambridge: Cambridge University Press, 2001.

Kemp, Ryan. “Kant and the Problem of Moral Conversion.” (Unpublished).

Konhardt, Klaus. “Die Unbegreiflichkeit der Freiheit. Überlegungen zu Kants Lehre vom Bösen.” *Zeitschrift für philosophische Forschung* 42 (1988): 397–416.

Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.

Langton, Rae. *Kantian Humility*. New York: Oxford University Press, 1998.

Loncar, Samuel. “Converting the Kantian Self.” *Kant-Studien* 104 (2013): 346–66

McCarty, Richard. “Kantian Moral Motivation and the Feeling of Respect.” *Journal of the History of Philosophy* 31 (1993).

———. *Kant’s Theory of Action*. New York: Oxford University Press, 2009.

Nauckhoff, Josefine. “Incentives and Interests in Kant’s Moral Psychology” *History of Philosophy Quarterly* 20 (2003): 41–60.

O’Neill, Onora. *Acting on Principle*. Cambridge: Cambridge University Press, 1975.

———. *Constructions of Reason*. Cambridge: Cambridge University Press, 1989.

Paton, H.J. *The Categorical Imperative*. Philadelphia, PA: University of Pennsylvania Press, 1971.

Pereboom, Derk. “Kant on Transcendental Freedom. *Philosophy and Phenomenological Research* 73 (2006): 537–67.

Reath, Andrews. *Agency and Autonomy in Kant’s Moral Theory*. New York: Oxford University Press, 2006.

———. “Kant’s Theory of Moral Sensibility.” *Kant-Studien* 80 (1989): 284–302.

Schafer, Karl. “Practical Cognition and Things-in-Themselves” in *The Idea of Freedom: New Essays on the Interpretation and Significance of Kant’s Theory of Freedom*. Edited by Dai Heide and Evan Tiffany. Oxford: Oxford University Press (forthcoming).

Stang, Nicholas. “The Non-Identity of Appearances and Things in Themselves.” *Noûs* 47 (2013): 106–36.

Strawson, P.F. *Bounds of Sense*. New York: Routledge, 1966.

———.

Sweet, Kristi. *Kant on Practical Life: From Duty to History*. Cambridge: Cambridge University Press, 2015.

Tenenbaum, Sergio. “The Idea of Freedom and Moral Cognition in *Groundwork* III.” *Philosophy and Phenomenological Research* 84 (2012): 555–89.

Ware, Owen. “Kant on Moral Sensibility and Moral Motivation.” *Journal of the History of Philosophy* 52 (2014): 727–46.

———. “Rethinking Kant’s Fact of Reason.” *Philosophers’ Imprint* 14 (2014): 1–21.

Watkins, Eric. *Kant and the Metaphysics of Causality*. Cambridge: Cambridge University Press, 2005.

Watkins, Eric and Willaschek, Marcus. “Kant’s Account of Cognition.” *Journal of the History of Philosophy* 55 (2017): 83–112.

Wolff, Robert. *The Autonomy of Reason*. New York: Harper & Row, 1973.

Wood, Allen. *Kantian Ethics*. Cambridge: Cambridge University Press, 2008.

———. “Kant’s Compatibilism” in *Self and Nature in Kant's Philosophy* edited by Allen Wood, 73–101. Ithaca, NY: Cornell University Press, 1984.

1. See the Cambridge collection *Practical Philosophy*, 129. Cf. Mary J. Gregor’s editorial note about the inclusion of Kraus’s review in the Academy edition of Kant’s work, *Practical Philosophy*, 121. [↑](#footnote-ref-1)
2. Henry Allison, “Empirical and Intelligible Character in the *Critique of Pure Reason*,” 1. Of course, Allison is an ardent defender of Kant’s view of freedom. The context here is that he is responding to critics who think the view is incoherent. Similarly, Graham Bird says that Kant’s account is “widely regarded as unacceptable or incoherent” (Bird, *The Revolutionary Kant*, 696). [↑](#footnote-ref-2)
3. Onora O’Neill, *Constructions of Reason*, ix. The same is true of O’Neill. She also defends Kant’s view and argues for a two-standpoint interpretation. She is noting the scholarly ambivalence concerning Kant’s view: “Kant is revered for his unswerving defense of human freedom and respect for persons, and for his insistence that reason can guide action. He is also reviled for giving a metaphysically preposterous account of the basis of freedom” (ix). [↑](#footnote-ref-3)
4. Jonathan Bennett, “Kant’s Theory of Freedom,” 102–12. [↑](#footnote-ref-4)
5. Ralph Walker, *Kant*, *the Arguments of the Philosophers*, 148. Cf. Jonathan Bennett, *Kant’s Dialectic*, 210–14. [↑](#footnote-ref-5)
6. See, for example, Christine Korsgaard, *Creating the Kingdom of Ends*, ch. 6. Cf. Hilary Bok, *Freedom and Responsibility*. Some, like Bok, defend this more as a “Kantian” view and less as an interpretation of Kant. [↑](#footnote-ref-6)
7. See, for example, Eric Watkins, *Kant and the Metaphysics of Causality*, 343–61. [↑](#footnote-ref-7)
8. There are other ideas (like our idea of God), which Kant considers to be within the domain of both theoretical and practical reason, but freedom is given special treatment as its epistemic status is unique. See *KpV* 5:4. [↑](#footnote-ref-8)
9. Quoted above, this is how Kant describes freedom in the preface of the Second *Critique*. See *KpV* 5:3­–4. [↑](#footnote-ref-9)
10. There are several places where Kant claims that transcendental freedom is an essential component of his moral philosophy. See, for example, *KpV* 5:93–106. [↑](#footnote-ref-10)
11. Although it will become clear in what follows that Kant has a technical and precise notion of the “incomprehensibility” of freedom, it is worth noting at the outset that this is importantly distinct from the colloquial sense of ‘incomprehensible’ which can mean something like ‘nonsense’ or ‘incoherent.’ [↑](#footnote-ref-11)
12. In-depth discussions of the incomprehensibility of freedom are rare. It is often mentioned only in passing as if it were something trivial or anodyne, not worthy of further investigation. I think this is a mistake, as I believe there are several interesting lessons to be learned from a close analysis of the issue. To the best of my knowledge, the only work dedicated entirely to the incomprehensibility of freedom is Klaus Konhardt, “Die Unbegreiflichkeit der Freiheit. Überlegungen zu Kants Lehre vom Bösen.” But the scope of Konhardt’s paper is largely restricted to the concept of evil, which is only one of the components of the incomprehensibility thesis. Another noteworthy exception is Allen Wood’s work, which shows a deep appreciation of the issue. See Wood, *Kantian Ethics*, 137 and “Kant’s Compatibilism,” 99. I worry that Wood might take it too far, however, as he treats Kant’s talk of transcendental freedom as total fiction. I explain how my view diverges from his in section 4. [↑](#footnote-ref-12)
13. The basic claim of the incomprehensibility thesis should be familiar to most Kant scholars; the gap in the literature, as I see it, is that the meaning behind this conclusion, the argument for it, and its many implications have been neglected. [↑](#footnote-ref-13)
14. See A445–51/ B473–79; A531–58/ B559–86 [↑](#footnote-ref-14)
15. Twenty might be a conservative number. It can be found in the following places: *G* 4:456, 4:459, 4:460, 4:461, 4:462; *KpV* 5:7, 5:45–49, 5:72, 5:94, 5:99, 5:133, *MS* 6:418; *KrV* A448/B476, A557/B585, *Rel.* 6:138, 6:145, 6:171, 6:191; and *KU* 5:275. It also pops up in various *Reflexionen* dating back as far the 1760s, and it appears in both metaphysics and ethics lectures. I quote some of these below, and I have numbered all the quotations containing the incomprehensibility thesis. [↑](#footnote-ref-15)
16. In *Kant’s Theory of Mind*, Ameriks puts this nicely: “Kant's main concern here was generally with the fact that we don’t and can’t have a theoretical proof of absolute freedom in the sense of an explanation of how such freedom works, that is, a speculative deduction or clarification of its noumenal operation” (195). Ameriks draws out the connection between the incomprehensibility of freedom and Kant’s claim that human beings are limited by the fact that we can explain events only in terms of laws of experience: “[F]or we happen to be beings who must always explain events by laws of experience, and these obviously are not to be found here” (*ibid*.). [↑](#footnote-ref-16)
17. Kant is especially clear about delineating thought and cognition in the second edition. See B146 and Bxxvi. [↑](#footnote-ref-17)
18. Describing the intelligible self as the bare idea that remains after we exclude all content from the world of sense is strikingly similar to what Kant says about the “negative” concept of noumena. See the “Phenomena and Noumena” section of the first *Critique* (A235–260/ B294–315). [↑](#footnote-ref-18)
19. The repetition of the word ‘comprehend’ throughout the paper is burdensome but necessary. ‘Comprehend’ is a technical term for Kant, and he uses it consistently in the context of freedom (he gives both the German ‘*begreifen*’ and the Latin ‘*comprehendere*’). In the *Jäsche Logic*, he distinguishes between seven grades of cognition, and comprehension is the highest. For a detailed discussion of cognition, including Kant’s seven-fold distinction in the *Jäsche Logic*, see Watkins and Willaschek, “Kant’s Account of Cognition.”

    Most important for my purposes in this paper: “The seventh, finally: to comprehend [*begreifen*] something (*comprehendere*), i.e., to cognize something through reason or a priori to the degree that is sufficient for our purpose. For all our comprehension is only relative, i.e., sufficient for a certain purpose; we do not comprehend anything without qualification” (LJ 9:64–65). In the *Vienna Logic*, Kant provides the example of our ability to comprehend the path of the moon sufficiently to predict a lunar eclipse (24:846). We can comprehend the moon’s orbit and make such predictions through reason without having to wait and see whether or not an eclipse will occur on a particular day.

    Thus, when Kant says that “we will never comprehend how freedom is possible [*wie Freiheit möglich sei, niemals begreifen Könnte]”* (*G* 4:456) or when he refers to the “incomprehensibility [*Unbegreiflichkeit*]” (*KpV* 5:7) of freedom, this technical definition is what he has in mind. We have no insight a priori through reason into freedom which is sufficient for theoretical purposes. As I argue below in the context of the **theoretical knowledge constraint**, our comprehension of freedom is perfectlysufficient for practical purposes but not theoretical purposes. See **(4)**, **(5)**, **(6)** and *Rel* 6:145. I elaborate on this below. [↑](#footnote-ref-19)
20. As I explain below, failing to have determinate content given through intuition is only one of two reasons why we cannot have theoretical cognition of a spontaneous first cause. The second reason is that our empirical, schematized concept of causality, as Kant explains it in the Second Analogy, would not able to process such a cause. [↑](#footnote-ref-20)
21. In this section, I identify four central themes of the incomprehensibility thesis; I call two of them “constraints” and the other two “problems.” I think of the problems as components of transcendental freedom that we cannot comprehend. The “constraints” by contrast are foreclosed possibilities of comprehending freedom. They pertain to means of cognitive access, which we do not possess, that would allow us to comprehend freedom. Although I treat these separately, they are closely related points. It could even be argued that the **cognition constraint** already comprises the other three components or that they are downstream effects of this constraint. Although I think that much of the issue comes back to the fact that we have no theoretical cognition of things in themselves, I have treated the four components separately because they speak to different aspects of the incomprehensibility of freedom. [↑](#footnote-ref-21)
22. Kant provides definitions for these terms in a few different places. In the *Groundwork*, Kant says that an incentive is “the subjective ground of desire” (*G* 4:428). In the second *Critique*, he says that an incentive is “the subjective determining ground of the will of a being whose reason does not by its nature necessarily conform with the objective law” (5:72). And he says that an interest “signifies an *incentive* of the will insofar as it is *represented by reason*” (*KpV* 5:79). Cf. *MM* 6:212, *G* 4:414n. In many cases, empirical incentives serve as the ground of our interest. But that does not mean that incentives alone suffice for determining the will. Reason must play a role. Many commentators have stressed the importance of Kant’s claim in the *Religion* that a human being’s power of choice (*Willkür*) “cannot be determined to action through any incentive except so far as the human being has incorporated it into his maxim” (*Rel* 6:24). See Allison, *Kant’s Theory of Freedom*, 51–52. For more on incentives, interests, and Kant’s moral psychology, see Josefine Nauckhoff, “Incentives and Interests in Kant’s Moral Psychology”; Patrick Frierson, *Kant’s Empirical Psychology*; Stephen Engstrom, “The *Triebfeder* of Pure Practical Reason”; and Herrera, “Moral Triebfeder.” [↑](#footnote-ref-22)
23. Kant reiterates this point in nearly identical terms in the *Critique of Practical Reason*: **(4)** “For, how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem and identical with that of how a free will is possible” (5:72). [↑](#footnote-ref-23)
24. The “fact of reason” argument is infamous for at least two reasons. First, it is generally considered to be a bad argument. Some think of it as question-begging at best, a “moralistic bluster” or “a spectacular train wreck” at worst. See Guyer, “Naturalistic and Transcendental Moments”, 462 and Wood, “*Kantian Ethics*, 135. For pushback against this consensus and an attempt to rehabilitate the argument, see Owen Ware, “Rethinking Kant’s Fact of Reason.” See also Karl Ameriks’s reconstruction of the fact of reason argument in his essay, “Is Practical Justification in Kant Ultimately Dogmatic?” in *Kant on Practical Justification*. Cf. Sasa Josifovic and Jörg Noller’s forthcoming *Freiheit Nach Kant*.

    Second, the fact of reason argument looks like a significant departure from *Groundwork* III. In the *Critique of Practical Reason*, Kant argues that we cannot start with freedom and then proceed to the moral law (5:29), but that is precisely what he seems to do in the *Groundwork*. Karl Ameriks has called this “a great reversal” and many (perhaps most) commentators have accepted his view. See Ameriks, “Kant’s Deduction of Freedom and Morality” and *Kant’s Theory of Mind*, 226. Nevertheless, there are some who argue that this is not in fact a reversal and that the two texts can be reconciled. For a recent argument to this effect, see Sergio Tennenbaum “The Idea of Freedom and Moral Cognition in *Groundwork* III.” For earlier work in favor of the “reconciliation” view, see Paton, *The Categorical Imperative*; Dieter Henrich, “Die Deduktion des Sittengesetzes,” and Lewis White Beck, *Commentary on Kant’s* Critique of Practical Reason. [↑](#footnote-ref-24)
25. One of his strongest statements of this constraint comes from *Religion within the Boundaries of Mere Reason*: “Hence in a practical context (whenever duty is at issue), we understand perfectly well what freedom is; for theoretical purposes, however, as regards the causality of freedom (and equally its nature) we cannot even formulate without contradiction the wish to understand it” (*Rel* 6:145). Cf. *Metaphysik Mongrovius* 29:918.

    I am inclined to go against the grain and call this practical “knowledge,” siding with Patrick Kain, Karl Schafer, and others. And (to some extent) this would put me at odds with Andrew Chignell. Chignell divides subjective grounds for assent (*Fürwahrhalten*) into epistemic and non-epistemic grounds, and, since we do not have theoretical reasons for believing that we are free, he lumps freedom in the category of “moral belief” along with belief in God and the afterlife. Unlike knowledge, practical belief (*Glaube*) does not have an objectively sufficient ground, but moral belief is also distinct from other forms of assent — persuasion (*Überredung*) and opinion (*Meinung*) — because moral beliefs have a non-epistemic (*viz*., practical) subjective ground. See Chignell, “Belief in Kant.” Cf. Kain, “Practical Cognition, Intuition, and the Fact of Reason,” 217–25 and Schafer, “Practical Cognition and Knowledge of Things in Themselves.”

    I must concede, however, that Chignell is right that the strict grounds that are required for *Wissen* are not given in the case of freedom, and although Kant occasionally refers to “practical cognition” (*praktischen Erkenntnis* [e.g., B xxi; *KpV* 5:20, 5:57, and 5:103]), he never uses the words ‘*praktisches* *Wissen*.’ So our rational belief in freedom is akin to knowledge, but it might be appropriate to withhold the ‘knowledge’ label, given Kant’s technical definition of the term. [↑](#footnote-ref-25)
26. Quoting all the relevant passages would involve a great deal of redundancy; I have limited myself to some characteristic excerpts from these later works. In the *Critique of Judgement*, Kant writes **(11)**: “There is thus an unlimited but also inaccessible field for our faculty of cognition as a whole, namely the field of the supersensible, in which we find no territory for ourselves, and thus cannot have on it a domain for theoretical cognition either for the concepts of the understanding or for those of reason, a field that we must certainly occupy with ideas for the sake of the theoretical as well as the practical use of reason, butfor which, in relation to the laws from the concept of freedom, we can provide nothing but a practical reality, through which, accordingly, our theoretical cognition is not in the least extended to the supersensible” (*KU* 5:175). Cf. *KU* 5:195, and *KU* 5:275. [↑](#footnote-ref-26)
27. He writes **(12)**: “When a human being is conscious of a duty to himself, he views himself, as the subject of duty, under two attributes: first as a *sensible being,* that is, as a human being (a member of one of the animal species), and secondly as an *intelligible being* (not merely as a being that has reason, since reason as a theoretical faculty could well be an attribute of a living corporeal being). The senses cannot attain this latter aspect of a human being; it can be cognized only in morally practical relations, where the incomprehensible property of *freedom* is revealed by the influence of reason on the inner lawgiving will” (*MS* 6:418). [↑](#footnote-ref-27)
28. It comes up frequently in *Religion within the Boundaries of Mere Reason* **(13)**: “Thus freedom — a property which is made manifest to the human being through the determination of his power of choice by the unconditional moral law — is no mystery, since cognition of it can be *communicated* to everyone; the ground of this property, which is inscrutable [*unerforschliche*] to us, is however a mystery, since it is *not given* to us in cognition” (*Rel* 6:138). **(14)** “Hence in a practical context (whenever duty is at issue), we understand perfectly well what freedom is; for theoretical purposes, however, as regards the causality of freedom (and equally its nature) we cannot even formulate without contradiction the wish to understand it” (*Rel* 6:145). See also 6:59, 6:117–18, 6:171, and 6:191. [↑](#footnote-ref-28)
29. More precisely, the soundness of the main argument requires the soundness of only one auxiliary argument. Since both (a) and (b) are necessary conditions, if either one of them is not met, then we cannot comprehend transcendental freedom. Nevertheless, I believe both auxiliary arguments are sound and thus we have two distinct explanations for the incomprehensibility of freedom. [↑](#footnote-ref-29)
30. Reason’s demand for completeness manifests as an endless search for “unconditioned conditions.” This is one of the principal themes of the Transcendental Dialectic. For a helpful explanation of how reason’s demand for completeness goes awry and gives rise to ceaseless illusions, see Michelle Grier, *Transcendental Illusion*. [↑](#footnote-ref-30)
31. Eric Watkins emphasizes Kant’s claim that the noumenal world contains the grounds of the world of appearance “*and so too of its laws*” (*G* 4:453). He argues that Kant’s view is that free noumenal choices could give rise to the laws of nature in such a way that actions in the empirical world are caused by both a free choice and deterministic laws. See Watkins, *Kant and the Metaphysics of Causality*, 328ff. Derk Pereboom lends Kant a hand by way of Luis de Molina, suggesting that God could create all and only those transcendentally free agents whose choices “conform to the deterministic laws that God intends for the phenomenal world” (“Transcendental Freedom,” 557). Allen Wood suggests (but does not ultimately endorse) the view that timeless noumenal choices could somehow narrow the set of possible worlds leading to the existence of an empirical character that was freely chosen by my noumenal self. Jeanine Grenberg makes use of Kant’s “schematism” chapter, proposing that we can understand empirical character as the “schema” of the intelligible character See Grenberg, “In Search of the Phenomenal Face of Freedom.” [↑](#footnote-ref-31)
32. Wood, *Kantian Ethics*, 137. [↑](#footnote-ref-32)
33. These names for the views come from Richard McCarty, “Kantian Moral Motivation and the Feeling of Respect.” McCarty, who argues for the affectivist view, is somewhat responsible for sparking the debate. On the intellectualist side, Andrews Reath could be seen as the modern standard-bearer (although the intellectualist view was considered the orthodox view for many years). Reath’s, “Kant’s Theory of Moral Sensibility” came out shortly before McCarty’s, and a flurry of articles have been published in the wake of these two influential papers. Reath’s article was reprinted in *Agency and Autonomy in Kant’s Moral Theory* and McCarty’s can be found in his book *Kant’s Theory of Action*. In favor of the affectivist view, see Ameriks, “Problem of Moral Motivation”; Herrera, “Moral *Triebfeder*”; Grenberg, “Kant’s Theory of Action”; and Guyer, *Freedom, Law, and Happiness*. On the intellectualist side, see Henry Allison, *Kant’s Theory of Freedom*; Robert Wolff, *The Autonomy of Reason*; and O’Neill, *Acting on Principle*.

    There have also been some attempts to find a middle way between the viewssuch as Erica A. Holberg, “The Importance of Pleasure in the Moral for Kant’s Ethics.” Owen Ware recently argued for a position that is sympathetic to the affectivist stance but which avoids the “conflict of forces” view, which he and others rightly see as a problem for the affectivists. See Ware, “Kant on Moral Sensibility and Moral Motivation.” [↑](#footnote-ref-33)
34. As McCarty points out, some commentators take this to be a damning criticism of Kant’s view (“Kantian Moral Motivation, 422). It would certainly run against the grain of common sense if Kant is claiming that a mere intellectual judgment that we are morally obligated to do something is sufficient for us to do it. To be fair to the intellectualists, however, this is an oversimplified description of the view. Reath’s explanation of it is more nuanced. See Reath, *Agency and Autonomy*, 8–28. [↑](#footnote-ref-34)
35. The “philosophers’ stone” quotation is often mentioned in the debate, and I think it does not weigh heavily in either direction. I cite it here simply for its relevance to the incomprehensibility of freedom. [↑](#footnote-ref-35)
36. But the fact that this is incomprehensible does not necessarily mean that it is impossible. This point was raised above in quote **(9**): “But because the possibility of such freedom cannot be comprehended, it does not yet follow from this that, because we cannot comprehend it, there also could not be any freedom” (*Metaphysik Pölitz*; 28:270). [↑](#footnote-ref-36)
37. These were evident in **(2)** and **(3)**. [↑](#footnote-ref-37)
38. This charitable assumption can be hard to maintain at times as commentators sometimes make it sound like this is indeed what they are trying to do. For instance, in a paper on this topic of the moral incentive, Stephen Engstrom writes, “In his chapter on the springs of pure practical reason, Kant undertakes to explicate, in the light of the *Factum* of reason, *how pure reason is practical* in the case of the human being, and more generally in that of a finite subject having a share in this *Factum*” (“The *Triebfeder* of Pure Practical Reason,” 93; emphasis added). By my reckoning, it is a mistake to say that Kant’s aim in the second *Critique* is to explain “how pure reason is practical” as Kant says many times (e.g., **[2]**, **[3]**, and **[4]**) that we will never comprehend how pure reason is practical. To continue maintaining the principle of charity (which I think we should, given Engstrom’s careful textual analysis), it is likely that this is simply an unfortunate quotation; I have no doubt that he is aware of Kant’s claim that we cannot know how pure reason is practical. My contention in this section is largely that those involved in this discussion could do more to situate it clearly in terms of comprehensible moral phenomenology downstream of the incomprehensibility issue. [↑](#footnote-ref-38)
39. Although references to the incomprehensibility of freedom are scarce in this debate, it has not been completely ignored. Richard McCarty cites part of **(4) —** the beginning of Chapter III of the second *Critique* — where Kant says that explaining “how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem and identical with that of how a free will is possible” (*KpV* 5:72). Rather than attempting to solve this “insoluble problem,” Kant says that the task of the chapter is not to show “the ground from which the moral law in itself supplies an incentive but rather what it effects (or, to put it better, must effect) in the mind insofar as it is an incentive” (5:72). McCarty says that his analysis of the phenomenology of respect must therefore avoid the determining ground question (which is insoluble) and focus instead on the psychological effect that the moral law has. He says, “This psychological effect is the motivating feeling of respect for the law. It constitutes the subjective, force-aspect of motivation” (*Kant’s Theory of Action*, 177–78). [↑](#footnote-ref-39)
40. See *Rel* 6:29ff. [↑](#footnote-ref-40)
41. On the *Gessinung* view, see *Rel* 6:25. And as for frailty being a kind of evil (and thus pertaining to the noumenal realm), see *Rel* 6:32. [↑](#footnote-ref-41)
42. Patrick Frierson, in his illuminating account of Kant’s empirical psychology, mentions this limitation specifically: “Commentators are correct to note that Kant’s account of human agency cannot be limited to any empirical account; for Kant, the moral importance of human choice emerges only when humans’ empirical character is explained by reference to a free intelligible character as its ground. Thus however detailed an empirical account of human frailty, the evil of such frailty must ultimately be explicable in terms of a practical perspective on choices that does not reduce them to empirical causes but ascribes them to freedom” (*Kant’s Empirical Psychology*, 235). [↑](#footnote-ref-42)
43. Even putting aside the problem of an evil tree bearing good fruit, the incomprehensibility of atemporal agency is an issue here. The idea of conversion is inextricably tied to that of change. It requires temporal succession: first the agent was bad, then she became good. And Kant acknowledges as much: “But does not the thesis of the innate corruption of the human being with respect to all that is good stand in direct opposition to this restoration through one’s own effort? Of course it does, so far as the comprehensibility of, i.e. our *insight* into, its possibility is concerned, or, for that matter, the possibility of anything that must be represented as an event in time (change)” (*Rel* 6:50).

    Kant recognizes that we necessarily think of all changes as taking place in time and therefore in accordance with the causal necessity of nature. In a footnote on the same page he calls the freedom involved here “an inscrutable property” and he says yet again that we will never be able to comprehend the relation between freedom and determinism. He bails himself out once more by making use of the “ought implies can” principle. Given my view on the incomprehensibility of freedom, I agree with Guyer that the “ought implies can” bailout is the best move Kant can make when faced with such a problem. See Guyer, “Problems with Freedom,” 201–2. [↑](#footnote-ref-43)
44. Loncar, “Converting the Kantian Self” 355–58. [↑](#footnote-ref-44)
45. *Rel* 6:44 [↑](#footnote-ref-45)
46. Loncar, “Converting the Kantian Self,” 364. [↑](#footnote-ref-46)
47. *Ibid.*  [↑](#footnote-ref-47)
48. I am grateful to Ryan Kemp for this response to Loncar. See Kemp, “Kant and the Problem of Moral Conversion.” [↑](#footnote-ref-48)
49. There is a divide in the one-world camp between epistemological and metaphysical views. Both sides share the belief that Kant asserts only one set of objects, but they differ on how to interpret the transcendental distinction. According to the epistemological view, Kant’s distinction is between two ways of considering objects or two standpoints we could adopt. Versions of this view have been defended by Henry Allison, Graham Bird, Christine Korsgaard, etc. It would be impossible to give an exhaustive list of citations. According to the metaphysical interpretation, the distinction is ontological. For instance, Rae Langton argues that it consists in distinguishing between unknowable intrinsic properties and knowable extrinsic properties. See Langton, *Kantian Humility*. Cf. Lucy Allais, “Kant’s One World” and *Manifest Reality*. [↑](#footnote-ref-49)
50. This was the traditional reading for many years, going as far back as Jacobi. P.F. Strawson was critical of the view but he attributes something like it to Kant in *Bounds of Sense*, 238. See also Guyer, *Kant and the Claims of Knowledge*. Once again, I will cannot give a complete list of references. For a recent defense of the two-object view and a criticism of the metaphysical one-world interpretation, see Nicholas Stang, “The Non‐Identity of Appearances and Things in Themselves.” [↑](#footnote-ref-50)
51. The three quotes I give here involve (A), but (B) is not uncommon. For instance, Jonathan Bennet takes issue with the idea that his free noumenal choices might determine facts about the empirical world that predate his birth. See Bennett, “Kant’s Theory of Freedom.” Korsgaard makes a similar point: “Obviously, if we try to picture how Marilyn's freedom is related to the forces that determine her, we must imagine it either inserting itself somewhere into the historical process, or standing behind the laws of nature from which this historical process necessarily follows. And both of these pictures seem crazy” (*Creating the Kingdom of Ends*, 203). [↑](#footnote-ref-51)
52. Frierson, “Two Standpoints and the Problem of Moral Anthropology,” 84; emphasis added. I would agree with Frierson and Korsgaard that two-standpoint views do not have the difficult problem of explaining how the two worlds relate. The two-standpoint view claims that the incommensurability of the practical and theoretical standpoints allows us to avoid the issue entirely (although I share Eric Watkin’s concerns about this move [see *KMC*, 320ff]). My complaint here is that they should not consider this an advantage over the two-world view. They are right to avoid explaining the relation between the intelligible self and empirical self, but they are wrong to think that two-world views should commit to such an explanation. [↑](#footnote-ref-52)
53. Dana Nelkin, “Two Standpoints,” 565. [↑](#footnote-ref-53)
54. Korsgaard, *Kingdom of Ends*, 204. [↑](#footnote-ref-54)
55. Once again, I think Allen Wood is one of the few commentators whose work shows a deep appreciation of this point: “So suppose we persist in asking: ‘Just how do freedom and natural causality really relate to each other? What is the metaphysical truth about how they fit together without contradiction? Is the noumenal realm involved in that in any way?’ The only permissible Kantian reply to these questions is: ‘I do not know, and neither do you, and neither can anyone ever know anything about this’” (Wood, *Kantian Ethics*, 138). [↑](#footnote-ref-55)
56. Furthermore, advocates of the two-standpoint view seem to be arguing that freedom is incomprehensible (in the pejorative sense) on the two-world view, but it would surely be a mistake to think that one-world views are capable of making freedom entirely comprehensible. Many aspects of transcendental freedom would remain incomprehensible on any view. Given Kant’s view of morality, nothing can explain how it possible for us to take an interest in the moral law. I am indebted to an anonymous referee for highlighting the importance of this point. [↑](#footnote-ref-56)
57. Some commentators have appreciated this point. For example, in a recent paper, Ryan Kemp explores the various explanations of how conversion might be possible, shows why they fail, and concludes that Kant’s considered view is that conversion is a “mystery.” See Kemp, “Kant and the Problem of Moral Conversion.” I agree with Kemp’s “mystery” view, but, for the sake of clarity, I should mention that Kant puts forward a technical sense of “mystery” in the *Religion* when he discusses things like miracles, and freedom is not strictly “mysterious” in the same way that miracles are (because we have a law for freedom). I am grateful to an anonymous referee for bringing this to my attention. It would be fair to say, however, that freedom and conversion are indeed “mysterious” in the more general sense of the word. See, Ameriks, “Kant, Miracles, and *Religion*, Parts One and Two.” [↑](#footnote-ref-57)
58. Kant makes it clear in various places that his moral theory requires incomprehensible transcendental freedom. See, for example, *KpV* 5:97. There is a wonderful expression of this point in Kraus’s *Review of Ulrich’s* Eleutheriology: “[F]*reedom* insofar as it is the basis of morality *cannot be comprehended*, and insofar as it can be comprehended it cannot serve as the foundation of morality” (8:458). This essay was included in the Academy edition of Kant’s writings on the grounds that Kraus claimed that he “made use” of an essay from Kant when composing his review. Vaihinger published the review calling it “A Hitherto Unknown Essay by Kant on Freedom.” It can be found in the Cambridge collection, *Practical Philosophy*, 121–31. [↑](#footnote-ref-58)
59. Kant’s view of freedom has been called “incoherent,” “metaphysically preposterous,” “worthless,” and “a hopeless failure.” Supra note 1, 2, 3, and 4. [↑](#footnote-ref-59)
60. See A557–58 / B585–86. This kind of agnosticism about freedom is perfectly fitting for theoretical reason. One of Kant’s central aims in the first *Critique* was to reign in metaphysical speculation about things like God and freedom. Kant’s plan was to argue in the second *Critique* that practical considerations alone can give us reason to believe in these things that theoretical reason cannot know or comprehend. [↑](#footnote-ref-60)
61. The incomprehensibility thesis cannot shield Kant from all objections his concept of freedom, but it can certainly protect his view from certain kinds of objections. For instance, Jonathan Bennett called Kant’s view “worthless” on the grounds that it does not have a plausible theory of the relation between free choices and empirical character. But on my reading, that is a feature of Kant’s view, not a bug. See Bennett, “Kant’s Theory of Freedom,” 102–12. [↑](#footnote-ref-61)
62. It might be impossible to overstate the importance of freedom in Kant’s philosophy. Beyond the moral significance discussed in this section, in his political philosophy, Kant argues that freedom alone can legitimate or justify the existence of the state. It is “the only original right belonging to every man by virtue of his humanity” (*MS* 6:237). Anticipating Hegel, Kant suggests that the realization of freedom is the ultimate *telos* of all human history and culture See, e.g., *Conjectural Beginning of Human History*, 8:185. For a novel and illuminating defense of this idea see Kristi Sweet, *Kant on Practical Life: From Duty to History*. Kant even thinks that we are compelled to entertain the idea that all of nature is teleologically directed toward the human being — the only creature on the planet in possession of this supersensible faculty — as nature’s “final end.” See *Critique of the Power of Judgment* (5:434–47). Cf. *Teleological Principles*, 8:182. See also Guyer, *Kant’s System of Nature and Freedom*, ch. 11.

    Kant’s view is that without freedom, earth would be bereft of value, a meaningless rock floating through the universe: “If all creatures had a faculty of choice bound by sensuous impulses, the world would have no value; the inner value of the world, the *summum bonum*, is the freedom to act in accordance with a faculty of choice that is not necessitated. Freedom is therefore the inner value of the world” (27:1482). For further discussion, see Paul Guyer, *Kant on Freedom, Law, and Happiness*, 96. [↑](#footnote-ref-62)
63. Providing a full account of what practical cognition is and how it differs from theoretical cognition would require its own lengthy discussion. See Karl Schafer “Practical Cognition.” [↑](#footnote-ref-63)
64. I am grateful to Allen Wood for raising this objection and pressing me on this point. [↑](#footnote-ref-64)
65. This is the view Wood gives in *Kantian Ethics*, 137. As discussed above, Wood says that the story told in the resolution to the Third Antinomy shows the possible compatibility of freedom and determinism but it does not commit us to any of the entities or ideas it mentions. It should be treated like fiction. Similarly, Graham Bird highlights Kant’s talk of intelligible causes as providing a resolution to the antinomy even if they should turn out to be fictional “*erdichtet sein sollte*” (A545/ B573). See *Revolutionary Kant*, 697 and “Antinomies,” 238.

    On my view, it may be appropriate to regard transcendental freedom as akin to fiction in theoretical discussions like the Third Antinomy. Kant’s objective there was to show that it is logically possible that this kind of freedom could coexist with determinism, and such a claim involves no real commitment to the view. But that does not mean that practical reason must also regard transcendental freedom as mere fiction. [↑](#footnote-ref-65)
66. Wood’s suggestion is that we should make do with practical freedom. He thinks that transcendental freedom should be “quarantined from Kantian ethics just as strictly as if it carried the plague (*Kantian Ethics*, 138). But I cannot come up with any other way of rendering Kant’s remarks in the Preface to the second *Critique* other than as a firm commitment to the transcendental freedom that he outlined in the first *Critique*. [↑](#footnote-ref-66)
67. Supra n. 28. [↑](#footnote-ref-67)
68. Watkins, *Kant and the Metaphysics of Causality*, 328ff. [↑](#footnote-ref-68)
69. Pereboom, “Kant on Transcendental Freedom,” 557. [↑](#footnote-ref-69)
70. One important difference, of course, is that the caffeine case is one of lacking comprehension in practice; whereas transcendental freedom is incomprehensible in principle. [↑](#footnote-ref-70)
71. *Rel* 6:45–53. For a discussion of the movement here and Kant’s conclusion that conversion is a mystery, see Ryan Kemp, “Kant and the Problem of Moral Conversion.” [↑](#footnote-ref-71)
72. I cited it above as **(4)**; see n. 21. [↑](#footnote-ref-72)
73. The resemblance between respect and the feeling of the sublime is striking. Kant takes note of this similarity in §27 of the third *Critique*. See *KU* 5:257–60. [↑](#footnote-ref-73)
74. Acknowledgments removed [↑](#footnote-ref-74)