

ON THE TRANSCENDENTAL FREEDOM OF THE INTELLECT

COLIN MCLEAR

University of Nebraska–Lincoln

Kant holds that the applicability of the moral ‘ought’ depends on a kind of agent-causal freedom that is incompatible with the deterministic structure of phenomenal nature. I argue that Kant understands this determinism to threaten not just morality but the very possibility of our status as rational beings. Rational beings exemplify “cognitive control” in all of their actions, including not just rational willing and the formation of doxastic attitudes, but also more basic cognitive acts such as judging, conceptualizing, and synthesizing.

Considered merely as intelligence, a human being is an imputable subject in which, of course, at the same time is united a natural being that subjects it to natural laws; [a being] nevertheless, whose determination depends upon intelligence, insofar as the possibility to act is granted him; therefore it is only to be derived how far a human being can be called cause of his actions and these can be imputed to him; and [it is] just as certain that, if he were led merely by natural laws, it would be impossible to impute to him any action, since the ground of action then would never lie in his control, but rather would be determined in the previous time.

Metaphysik Vigilantius (K₃), 29:1020 (1794/95)
IMMANUEL KANT

1. Introduction

One well-known point of emphasis in the critical philosophy is that only transcendental idealism can safeguard the possibility of the spontaneity and agent-causal freedom of rational beings. If such freedom were not possible, then in Kant’s estimation there would be no hope for conceiving of rational agents as morally responsible; for if rational agents were totally in the grip of the deterministic

causal nexus of the spatio-temporal world then moral requirements could not apply.

However, as the epigraph above makes clear, Kant considers the causal nexus of phenomenal nature to threaten the status of human beings not just as moral agents but even as “intelligences”—that is, as beings capable of being determined in their acts by their intellectual faculties.^{1,2} Phenomenal causes, if they were the only possible causes upon which action was based, would threaten the very possibility of our status as rational beings. Hence, according to Kant, the rational acts of rational agents depend for their generation on an intellectual faculty whose acts are independent of the phenomenal causal nexus.

This paper has two primary aims. The first is to show that Kant’s position concerning the freedom of the intellect is a central commitment of the critical philosophy, entailed by core elements of his conception of the nature of the intellect and its activity.³ Because Kant’s conception of freedom is often linked more closely to his practical philosophy, and specifically his conception of a free *will*, or practical reason, I focus here primarily on the *theoretical* use of the intellect.

Second, by emphasizing Kant’s conception of the freedom of intellectual acts, I hope to widen the scope of discussion on this issue, which has largely focused

1. Here and throughout I use “intellect” to designate the “higher” faculty of cognition, in contrast to sensibility, and which includes at least the understanding, judgment, and reason. Kant at least sometimes uses “*Verstand*” in this general sense (e.g., A294/B350; Pr 4:288), in contrast to a narrower use where it designates the faculty through which concepts are generated by means of reflection. He also sometimes talks of “reason” (*Vernunft*) in a similarly general sense. I also follow Kant in sometimes speaking of practical and theoretical *uses* of the intellect.

2. Kant characterizes the spontaneous cognizing self as an “intelligence” at several points (e.g., B155, B157–58, B158n). He also sometimes identifies this cognizing self or intelligence as a noumenal entity (e.g., *Metaphysik K*₂ 28:773 (1790/1); see also *Metaphysik Mrongovius* 29:926 (1782/83); G 4:446ff; CPrR 5:114; *Metaphysik K*₂ 28:775; Refl. 23:34, 35 (CV E38)). There is some dispute as to how to understand Kant’s distinction between an “intelligence” and a “soul”. See, e.g., Allison (1996b: 65–66), Van Cleve (1999: 182–83), Wuerth (2014: 43, 63–64). As I understand him, Kant construes a soul as being or having a “principle of life,” which he construes as a faculty for acting on the basis of representations (e.g., *Metaphysik Dohna* 28:682 (1792/93), *Metaphysik K*₂ 28:755). Kant allows that non-human animals (i.e., irrational animals) are alive in this sense, while lacking an intelligence in his more substantive sense, which is coherent if he construes intelligences as living beings (souls) with intellectual faculties, while non-human animals are souls with only a faculty of sensible receptivity. See, e.g., *Metaphysik L*₁ 28:275 (1778–81), 276; *Metaphysik L*₂ 28:594 (c. 1790); *Metaphysik Dohna* 28:679, 690. For further discussion of life as an internal principle or determination of action, as opposed to an external physical determination see Frierson (2014: 54–56). Thanks to Yoon Choi for bringing many of these texts to my attention.

3. Perhaps the most prominent recent account congenial to my thesis is that of Henry Allison (1990), who likewise argues that the intellect must be free both in its theoretical and its practical use. Besides some differences of approach and emphasis there are, however, some important differences with regard to the substance of our views, which I discuss further at various points below. See especially Footnote 55.

on Kant's conception of the formation of doxastic attitudes—varieties of “holding-for-true” (*Fürwahrhalten*) or “assent”.⁴ I argue that accounts of Kant's conception of the freedom of the intellect that focus solely on doxastic attitude formation are overly narrow. Kant's account of attitude formation is inextricably linked to his broader conception of the freedom of rational or intellectual activity *as such*, and thus for better or worse, tied to other instances of such free activity, such as intellectual forms of synthesis, which are free even if not directly or immediately under one's voluntary control. I thus argue that Kant's conception of the control we have over doxastic attitude formation is rooted in a more general conception of *cognitive* control, which includes not just the formation of doxastic attitudes, but also more basic cognitive acts such as judging, conceptualizing, and synthesizing. There are, to be sure, important differences between these types of act, and thus between the different ways in which the freedom of rational agency manifests itself in finite beings. But these differences are part of an underlying unity, one which I attempt to bring out in the discussion below.

The core conception of cognitive control for which I advocate hinges on a proper understanding of Kant's conception of an intellectual “power/capacity” (*Fähigkeit*) or “faculty” (*Vermögen*)—I will largely use these terms interchangeably.⁵ One has cognitive control just in case one's acts are the result of the exercise of one's faculties according to the constitutive laws of their activity, without being determined by grounds whose existence and nature are independent of the capacity being exercised. Hence, in a sense laid out further below, beings in cognitive control of their acts must be such as to causally act without being causally determined by anything else. In Kant's terms they must thus be “absolutely spontaneous” or “transcendentally free.” This notion of cognitive control is extensionally equivalent to, but more conceptually basic than, Kant's conception of transcendental freedom, particularly in the manner in which it articulates the

4. For discussion of assent see Stevenson (2003), Chignell (2007a; 2007b), Pasternack (2011). For discussion of freedom in attitude formation see, e.g., Henrich (1975), Allison (1990; 1996a), Schönecker (1999), Bilgrami (2006), Boyle (2009), Kitcher (2011: ch. 14), Kohl (2015b). Why the discussion has been limited in this way is not altogether clear, but may be due to a conception of action solely in terms of voluntary action, of intellectual activity as non-voluntary, and thus of the intellect as an inappropriate subject of discussion with respect to freedom. For such a view see, e.g., Kitcher (1990: 122), which construes many of the kinds of intellectual act I shall discuss as “subpersonal” and thus not free in any relevant sense. Kohl (2015b: 309, note 18) distinguishes the (free) intellectual act of assent/judgment from the forms of synthesis involved in more basic intellectual acts. I argue that Kant's account of the freedom of the intellect should not be divided in this way. If Kant has a coherent conception of the transcendental freedom of the intellect at all, it is a unified one.

5. For discussion of why they ought to be distinguished, at least in some cases, see McCarty (2009: 18–20) and Smit (2009: 240–43). For my purposes here talk of ability, capacity, power, or faculty is largely interchangeable. What *is* important to distinguish is the difference between a capacity or faculty (*Vermögen, facultas*) and a force (*Kraft, vis*), especially when the latter term is sometimes translated as “power.” For further discussion see also Longuenesse (1998: 7–8).

source of the incompatibility of rational action with a temporally determined causality.

In Section 2 I characterize Kant's conception of an intellectual act and the related notion of an intellectual faculty or capacity. Section 3 describes Kant's conception of transcendental freedom as a form of causal ultimacy, its links with the above conception of control, and Kant's view of the sense in which free rational beings possess "leeway" or the ability to do otherwise. Section 4 then formulates and critically discusses four arguments elucidating the way in which Kant implicitly or explicitly regards the intellect as free. In Section 5 I discuss two general objections regarding first, whether Kant might not endorse a weaker and merely "relative" spontaneity for the theoretical use of the intellect, and second, whether or how his views on the freedom of the intellect change over the course of his critical writings. Finally, I summarize and conclude the paper.

2. Intellectual Acts

To say that acts of the intellect are free requires the coherence of speaking of intellectual "acts" at all. Kant construes the notion of an act (*Handlung*) as closely connected with the concepts of activity (*Tätigkeit*), substance, and force (*Kraft*) (A204/B250; cf. R5289–90 (1776–78?) 18:144, R5650 (1785–88) 18:298–302). Kant's critical conception of causality is also closely linked to these notions, for an act (*Handlung*) "already signifies the relation of the subject of causality to the effect" (A205/B250). In his metaphysics lectures from the critical period, Kant is reported to describe an act as the determination of substance as the cause of an accident (*Metaphysik Pölitz* 28:564–65 (1790/1)).⁶ Thus, as I shall use the notion in this paper, an act is an actualization of a substance—through one of its capacities or faculties—to exert a force that engenders some accident as an effect of that actualization.⁷ This conception is quite broad, and captures much more than merely bodily movements, as some commentators have tended to construe Kant's

6. For extensive discussion of this conception of an act, as well as the related notions of substance and power in Kant's work see Watkins (2005: chs. 4–5), Smit (2009), Wuerth (2014: especially chs. 1, 6), Stang (2019: 92–94).

7. For versions of this kind of approach see also Willaschek (1992: chs. 2, 11), Watkins (2010). There is an important question as to whether Kant conceives of the actualization of a capacity as (i) grounded in a law or of the law as (ii) grounded in the nature of the substance and its capacities, or whether the law and substantial nature are (iii) coeval. I remain neutral on this issue here. However, I will sometimes speak of a faculty as "governed" or "constituted" by its laws. Nevertheless, I intend this way of speaking to be neutral with respect to the grounding relation between a substance's nature and the laws related to the exercise of its causal powers. For further discussion of this issue and its Leibnizian background see Watkins (2005; 2019a: chs. 1–2, 11), Kreines (2008), Rutherford (2013), Stang (2016a: ch. 8), Kreines (2017), Messina (2017).

conception of action.⁸

One might worry that conceiving of acts in this way is in fact *too* broad. Perhaps we should take Kant as reserving “act” and its cognates for bodily behavior. However, Kant distinguishes “inner” and “outer” exercises of voluntary agency (MM 6:214, 218–19), which means he denies that agency per se requires (even typically) the exercise of one’s body.⁹ He also claims that, with respect to the practical exercise of one’s reason, the only acts of choice that are always fully within one’s control are inner acts (CPrR 5:36–37; CPJ 5:471n). Kant also refers to “actions of the understanding [*Verstandeshandlungen*]” (e.g., A330/B387) and to the exercise of one’s attention as a free and “real act [*Act*] of the faculty of cognition” (An 7:131).¹⁰

That Kant’s position on the nature of the acts of a substance, via an exercise of some or other of its capacities, concerns acts per se and not simply outer bodily actions or inner intentions is further supported by his discussion of the intellect’s acts in the Transcendental Dialectic’s explication of illusion.

while [the intellect] acts merely according to its own laws [*nach seinen Gesetzen handelt*], its effect [*die Wirkung*] (the judgment) must necessarily agree with these laws. (Transcendental Illusion, A294/B350)

The passage indicates that the intellect acts according to laws governing its effects, which include the judgments it forms. If acts are generically actualizations of powers to cause determinate effects according to laws, then intellectual acts are such acts as governed by rational or intellectual laws—laws that Kant sometimes describes as laws of “combination” (*Verbindung*):

combination is . . . not given through objects but can be executed only by the subject itself, since it is an act of its self-activity. (Transcendental Deduction, B130)

Intellectual acts occurring according to such laws of combination constitute a wide class. Amongst the class of intellectual acts that will concern us below are

8. Kohl (2015b: 314 note 34) raises this issue as well. Those who tend to distinguish thinking from acting include Henrich (1975: 66), Kosch (2006: 34), Kitcher (2011: 178); see also Allison (1990: 63, 218), where he consistently distinguishes between rational beings (i.e., thinkers) and rational *agents*. I see no such restriction in Kant’s usage. However, since Allison endorses a unified conception of rational agency in Kant, it isn’t clear whether this is really a substantive problem for him. For other positions close to the one I express here see Paton (1947: 209), Wood (1999: 172), Watkins (2005; 2010). Engstrom (2009) also emphasizes the importance of construing the activity of the cognitive faculties in causal terms.

9. Note that as I understand Kant, the genus of voluntary acts includes those that are, either immediately or mediately, determined by the agent’s will or power of choice. I provide a more detailed discussion of voluntary action in Section 4.1.

10. For related discussion see Kohl (2015b: §2), Merritt (2018: chs. 2–4).

acts of “assent” or holding-for-true (*Fürwahrhalten*), acts of thinking or judging, and other even more basic cognitive acts. Let me say a bit more, respectively, about these different classes of act.

Acts of assent are acts by which a subject comes to hold what we would now call a propositional attitude (specifically a doxastic attitude) towards some truth-functional content, on the basis of either or both “objective” and “subjective” grounds, and whose adoption is related to the subject’s assessment of the probability of the truth of the judgment.¹¹

Thinking is the act of “unifying representations together in one consciousness” (Pr 4:304; see also JL 9:101; *Wiener Logik* 24:928 (1780–82)). Thinking is opposed to merely associating, and as I will typically use it, concerns all acts of conceptualizing, judging, and inferring.

Kant also characterizes a third class of basic cognitive acts necessary for thought, or at least conceptualization and judgment, such as in his discussion of attention and abstraction where he says,

The endeavor to become conscious of one’s representations is either the *paying attention to (attentio)* or the *turning away from* an idea of which I am conscious (*abstractio*). – The latter is not the mere failure and omission of the former . . . but rather a real act of the cognitive faculty of stopping a representation of which I am conscious from being in connection with other representations in one consciousness. . . . To be able to abstract from a representation, even when the senses force it on a person, is a far greater faculty than that of paying attention to a representation, because it demonstrates a freedom of the faculty of thought and the authority of the mind, *in having the object of one’s representations under one’s control [Gewalt]*. (An 7:131–32)

Though Kant here distinguishes abstraction from attention, abstraction is itself better understood as a *form* of attention.¹² If attention is the focus on, or striving-to-bring-to-consciousness of, some representational content, abstraction is the capacity to focus on or attend to some subset of that content—to disregard some

11. See Kant’s discussion at A820/B848ff, Chignell (2007a; 2007b), and the references cited in Footnote 4 above. Note that the fact of a distinction between making a judgment and assenting to that judgment does not mean that acts of judgment can be made entirely independently of *any* attitude towards the truth of those judgments. Kant claims that all judgments have a “modality” (A74/B99–100). “Assertoric” and “apodictic” judgments will be accompanied by some form of assent. “Problematic” judgments are those the truth of which is not in any way endorsed or asserted. Kant claims that one can problematically judge even that which one knows is false (A75/B100–101). Thus the only constraint on problematic judgment is logical possibility (i.e., the laws of identity and non-contradiction). Thanks to an anonymous referee for encouraging clarity on this point.

12. See also Merritt and Valaris (2017: 576–80).

part of it in favor of another. Part of what distinguishes abstracting from attending *simpliciter* is that abstraction necessarily requires a degree of control over what is to be considered the object of representation whereas attention *simpliciter* does not. In the latter case one may be passively drawn to some representation (or the content thereof). In the case of abstraction one actively exerts control over the object of one's attention. Kant describes this control over one's representations as a manifestation of a form of cognitive freedom, as it requires that the subject can determine the object of representation in a manner that is independent of the nature of the original representational stimulus, cause, or occasion.

Hence, assent, thought, and correlated cognitive acts, such as attention and abstraction, are all going to be intellectual "acts" in the sense defined above. Such acts are acts that are not necessarily all voluntary (though some surely are, like directed acts of attention), but are nevertheless under the subject's control (in a sense to be further discussed). I argue below that, for Kant, a particular form of control is a necessary condition for the occurrence of such acts and that this form of control, properly understood, requires that intellectual acts must be transcendently free. In order that we understand what this means I turn now to a discussion of these notions of freedom and control.

3. Freedom & Control

In this section I first discuss two different kinds of freedom Kant distinguishes in his writings from the critical period.¹³ These are what he terms "practical" and "transcendental" freedom. I take these in turn. I then discuss the relationship between Kant's characterization of transcendental freedom as causal ultimacy and his notion of control. I argue for the view that the kind of cognitive control characteristic of intellectual activity entails that it must be transcendently free activity. Finally, I connect the issues of freedom and control with Kant's conception of the "leeway" possessed by an agent to do otherwise than she in fact does, and how this grounds "ought" claims or imperatives to which the agent is subject.

3.1. *Practical & Transcendental Freedom*

We can understand practical freedom as follows:

Practical Freedom: the capacity to act in accordance with ends distinct from those dictated by one's immediate sensible impulses.

Practical freedom is typically demonstrated by rational agents pursuing ends even when they conflict with the agent's immediate inclinations or interests.

13. For discussion and overview see Beck (1987), Carnois (1987), Allison (1990), Kosch (2006: ch. 1).

Rational agents are capable of pursuing ends in the face of contravening desires or competing pleasures. As Kant points out,

we have a capacity to overcome impressions on our sensory faculty of desire by representations of that which is useful or injurious even in a more remote way. (Canon, A802/B830)

The famous “marshmallow tests” examining the extent of a child’s capacity for delaying gratification, as conducted by Walter Mischel in the 60s and 70s, can be understood as testing the extent or strength of this capacity for practical freedom.¹⁴

While Kant is clearly talking, in the Canon quote above, about the capacity, independent of current competing desires and inclinations, to form intentions for carrying out bodily actions, much the same point can be made concerning intellectual acts of the kind discussed in the previous section. Rational adults exhibit, at least some of the time, an ability to conceptualize, judge, infer, and ultimately assent on the basis of something other than what the subject currently wants, desires, or otherwise has an immediate inclination to act on, and often contrary to what the immediate deliverances of sense would seem to indicate.¹⁵ In particular, such subjects can engage in intellectual acts that respect canons of evidence and laws of truth-preservation even if such canons and laws yield verdicts contravening one’s current interests, inclinations, or the strength and character of one’s sensations. Hence, the core notion of “practical freedom” is just as significant for theoretical uses of the intellect as it is for the more traditionally discussed practical uses exemplified by acts such as willing or intending.

Are there further forms of freedom that a subject must exhibit in order to possess practical freedom? In particular, does practical freedom require what Kant calls “transcendental freedom”? Kant discusses the notion of transcendental freedom in the resolution of the Third Antinomy. He says,

By freedom in the cosmological sense, on the contrary, I understand the power of beginning a state **from itself** [*von selbst*], the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature. . . . It is especially noteworthy that it is this **transcendental** idea of **freedom** on which the practical concept of freedom is grounded, and the former constitutes the real moment of the difficulties in the latter, which have long surrounded the question of its possibility. (A533/B561; see also A445/B473 and A446/B474)

14. See Mischel, Ebbesen, and Raskoff Zeiss (1972), though see Watts, Duncan, and Quan (2018) for critical discussion.

15. I discuss the issue of whether judgment or assent is under one’s voluntary control in the next section.

So, for Kant, the idea of freedom, understood as “transcendental freedom”, is the idea of an “absolutely spontaneous” act (A448/B476)—that is, an act whose ultimate causal source lies in the subject, and is not itself determined by any temporally structured causal ground.¹⁶ Kant denies that we can ever know whether an act is free in this sense (A557/B585). Kant also denies here that practical freedom is wholly independent of transcendental freedom; the practical conception is “grounded” on the transcendental, and he subsequently says that “the abolition of transcendental freedom would also simultaneously eliminate all practical freedom” (A534/B562). This seems to be Kant’s considered view, for in the discussion of freedom in the second *Critique* Kant also claims that without transcendental freedom “no moral law is possible and no imputation in accordance with it” (5:97).

However, the conception of transcendental freedom so presented can at least *seem* to conflict with the discussion of freedom in the Canon of the first *Critique*, for Kant says in that late section that “Practical freedom can be proved through experience” (A802/B830), in that normal adult human beings clearly have practical freedom in the sense I defined above. Kant’s claim in the Canon presents a straightforward problem, for if practical freedom depends on absolute spontaneity/transcendental freedom, and we cannot know we have the latter, then we cannot know we have the former, and thus cannot “prove it” through experience.¹⁷

What to make of the text of the Canon then?¹⁸ One way to try and resolve this plays off of Kant’s claim in the Canon that

The question of transcendental freedom is a matter for speculative knowledge only, and when we are dealing with the practical we can leave it aside as being an issue with which we have no concern. (A803–804/B831–32)

One reason Kant might say that we can simply “leave aside” the issue of transcendental freedom when dealing with the practical, is that at time of writing the Canon Kant really did conceive of them as independent of one another. Hence, on this way of explaining the text, the conception of practical freedom in the Canon is an early statement of Kant’s view concerning the relation between

16. I treat “transcendental freedom” and “absolute spontaneity” as extensionally equivalent even if conceptually distinct—“transcendental freedom” connotes a feature of action while “absolute spontaneity” connotes a type of causal power; despite these subtleties I’ll tend to use the terms as basically interchangeable. See also Allison (1990: 15, 60); cf. Smit (2009: 242n) for a denial of their equivalence, though Smit does not fully explain why this may be. I discuss the issue of time and the problem it poses for free action further in Section 3.2 below.

17. Thanks to an anonymous referee for recommending this way of putting the problem, and for suggesting one possible solution.

18. This is sometimes called the “Canon problem”; see Schönecker (2005), Rosenkoetter (2019) for general discussion.

transcendental and practical freedom, one which he ultimately revises in the statement made in the discussion of the Third Antinomy (which was written after the text of the Canon, despite coming before it in the work), and which is subsequently given its clearest expression only in the second *Critique*.¹⁹

Another possibility is that Kant utilizes both a stronger and weaker sense of “practical freedom” in the Canon.²⁰ The weak sense is merely that of being able to resist immediate sensible impulses. This is something, as we’ve noted, of which normal adult humans are clearly capable, and is thus provable via experience. What is *not* provable through experience is possession of practical reason in the stronger sense—that is, that a causally ultimate faculty of *reason* explains the possession of the capacity for practical reason in its weak sense in normal adults. Kant may be indicating that it is this stronger sense that can be put aside when he says,

But whether in these actions [of resisting immediate sensible impulses], through which it prescribes laws, reason is not itself determined by further influences, and whether that which with respect to sensory impulses is called freedom might not in turn with regard to higher and more remote efficient causes be nature – in the practical sphere this does not concern us (Canon, A803/B831)

If it is merely nature that ultimately explains the weak sense of practical freedom, then we lack practical freedom in the strong sense, for we lack the capacity for reason to be causally ultimate in determining our wills. Hence, only practical freedom in this strong sense—as requiring determination by reason rather than nature—would depend on transcendental freedom.

In any case, whether Kant developed his view after writing the Canon into the more resolute view of freedom on display in the Third Antinomy and the second *Critique*, or he consistently maintained his view from the writing of the Canon, the result is the same—according to the mature critical philosophy rational determination of one’s actions presupposes transcendental freedom. I understand Kant’s conception of transcendental freedom as follows:

Transcendental Freedom: the power to (i) initiate a causal series from oneself (ii) without being determined to do so by any temporally structured causal ground.

A transcendently free rational agent must be the causally ultimate source of her actions—call this the “source requirement” for transcendental freedom. That

19. For discussion of this sort of “patchwork” explanation see Carnois (1987: 29), Allison (1990: 54; 1996b: 109ff), Kosch (2006: 19–21).

20. Thanks to an anonymous referee for suggesting this option.

is, the agent must be able to (i) initiate a causal series, which includes or constitutes her act, without being (ii) causally determined in this initiation by any temporally structured causal ground.²¹ Moreover, since Kant locates the source of transcendently free action in the agent's power to so initiate a causal series, it cannot be the case that transcendently free acts just *happen* as wholly random or indeterministically caused events in nature. As Kant notes of the view of Lucretius, such events or happenings would amount to "blind chance" (CPrR 5:95).

Hence, for Kant, transcendently free acts are not indeterministic, they are causally determined. It is just that they are causally determined *by the agent* and not by the mechanism of nature. This is particularly clear from Kant's discussion of "predeterminism" in the *Religion*,

[absolute spontaneity] is at risk only with predeterminism, where the determining ground of an action lies *in antecedent time*, so that the action is no longer in my control [*Gewalt*] but in the hands of nature, which determines me irresistibly. (Rel 6:49–50n)

Absolute spontaneity is not at risk from determinism but from *predeterminism*. Satisfaction of the source requirement thus depends on the agent's actions not being exhaustively causally determined by preceding events.²² Given the causal principle articulated in the Second Analogy, the natural world is such that, for any agent, the ultimate determinants of her actions would be in the past, and thus beyond her "control" (*Gewalt*).²³

For from the [necessity of causal relations] it follows that every event, and consequently every action that takes place at a point in time, is necessary under the condition of what was in the preceding time. Now, since time past is no longer within my control [*nicht mehr in meiner Gewalt ist*], every action that I perform must be necessary by determining grounds *that are not within my control*, that is, I am never free at the point of time in which I act. (CPrR 5:94; see also 5:95)

As Kant conceives of the causal order of the phenomenal world, it is such that within this order, one never has control over one's actions, and this lack of control is sufficient to undermine one's freedom. Transcendently free action within the phenomenal realm is thereby impossible (I discuss exactly why in the next

21. See also Wood (1984: 79–83), Pereboom (2006a: 542).

22. The distinction between "source" and "leeway" requirements for free action is common in the contemporary literature on free will. The terminology originates with McKenna (2001); see also Pereboom (2001; 2006a: 542). I discuss the issue of leeway further below.

23. See also Kosch (2006: 26), Pereboom (2006b: 539–41).

subsection). Transcendental freedom thus requires a different causal order, one whose possibility depends on features of negative noumena or things as they are in themselves.²⁴

It is worth emphasizing here that, as I understand Kant's conception of free action, it depends on the coherence of his distinction between things as they appear and things as they are in themselves, and thus on his transcendental idealism. The exact meaning of Kant's doctrine of transcendental idealism is, however, contested. I take it though that in determining the correct interpretation of Kant's idealism we are constrained by the conditions Kant sets on the kind of control outlined here—viz. the metaphysical conditions on the ultimate ground of explanation of a rational action. The correct interpretation must allow for a logically coherent distinction between a temporal and a non-temporal causality, with it being at least epistemically possible that rational agents possess such a non-temporal causal power. However, I leave open here whether, or to what extent, this conception of freedom is compatible with the wide array of readings of Kant's idealism, from purely epistemic “inference ticket” readings like that of Henry Allison, to more robustly metaphysical readings like those of Ameriks, Langton, and Allais.²⁵

3.2. *Transcendental Freedom & Control*

We've seen that Kant construes transcendental freedom in terms of a rational agent's causal ultimacy—what he calls the agent's “absolute spontaneity”—and that he thinks of the temporal causal structure of the phenomenal world as threatening such ultimacy by placing the causal grounds of action beyond the agent's control. But it is not entirely obvious *why* the temporal structure of phenomenal causation threatens the control necessary for transcendental freedom, or exactly how we should understand the notion of control at issue.

There seem to be at least three reasons why Kant rejects the compatibility of temporal causality with control—viz. that (i) it entails the existence of a causal ground distinct from and independent of the agent; (ii) it entails the existence of

24. Depending on how one understands Kant's transcendental idealism—e.g., as in a “one world” or “two world” way—one may have to modify the language I've used here—e.g., that with respect to its phenomenal aspect an action cannot be free, but might be with respect to its noumenal aspect, etc.

25. For representative discussion of transcendental idealism see Guyer (1987), Langton (1998), Allison (2004), Bird (2006), Ameriks (2012: chs. 3–5), Allais (2015), Stang (2016b). I am inclined to think that only a thoroughly metaphysically committal reading of transcendental idealism will serve Kant's position concerning the freedom of the intellect, but I lack the space to further pursue that claim here. For a worry about whether merely “epistemic” interpretations, like that of Allison's, are able to account for the activity of the intellect see Ameriks (1992: 215–18); for Allison's reply see Allison (1996b: 124–28). I discuss some of my differences with Allison's view in particular in Footnote 55 below.

a ground over which the agent can exert no causal influence; (iii) it entails the operation of a form of causality that is not characteristic of the activity of the agent's intellectual faculty. Kant does not always distinguish between these three reasons, but I think that while he often appeals to either (i) or (ii), in fact it is (iii) that is the root of the incompatibility between temporal causation and control, and which also explains his endorsement of (i) and (ii). Let me explain these in turn.

First, if the ultimate causal grounds of action can be traced to temporally remote causes that lie outside the agent, then it is clear why she would not be transcendently free—her act would not be initiated by her but by something else. Note here that it is not time *per se* that is the ground of her lack of freedom, but rather the fact that the ultimate cause was external to her. Time would be an aid to determinism but not itself directly undermining of the subject's control. Something like this issue seems to be at work in Kant's discussion of theological determinism and the transcendental reality of time in the *Critique of Practical Reason*. There he raises the worry that,

it . . . seems that, as soon as one admits that God as universal original being is the cause also of the existence of substance . . . one must admit that a human being's actions have their determining ground in something altogether beyond his control, namely in the causality of a supreme being which is distinct from him and upon which his own existence and the entire determination of his causality absolutely depend. In fact, if a human being's actions insofar as they belong to his determinations in time were not merely determinations of him as appearance but as a thing in itself, freedom could not be saved. A human being would be a marionette or an automaton, like Vaucason's, built and wound up by the supreme artist; self-consciousness would indeed make him a thinking automaton, but the consciousness of his own spontaneity, if taken for freedom, would be mere delusion inasmuch as it deserves to be called freedom only comparatively, because the proximate determining causes of its motion and a long series of their determining causes are indeed internal but the last and highest is found entirely in an alien hand. (CPrR 5:100–101)

There are a number of points made in the passage. Kant first suggests that if we are creatures of God then there seems to be a problem for freedom. Moreover, if time is transcendently real and God is the ultimate causal origin of the existence of all things in their complete determinacy then Kant further concludes that it would be unavoidable that all beings have their acts ultimately causally grounded in God. In acting on the basis of causal grounds determined by another, humans would thus be "marionettes," controlled by an external source.²⁶ This is even more

26. There is considerably more to the issue of theological determinism than I can engage

explicitly articulated in the lectures on ethics:

If we follow up the determining grounds of human actions [in time], they are linked to one another in a chain; if we go back to the source, the only possible outcome is that we must arrive at an external cause, a being that is outside the agent. (*Ethik Vigilantius* 27:505 (1793/4))

Hence, on this conception of the problem of (pre)determinism, time is merely its handmaiden. It is God's causally creative activity that provides the real threat to control, and the reality of time would make this threat unavoidable by making dependence on an external causal source for one's acts unavoidable.

However, there are (perhaps broadly Leibnizian or Spinozistic) conceptions of the causal structure of change in a subject's states that are temporal but do not (or need not) appeal to causally external factors. In such cases the causal ultimacy characteristic of transcendental freedom would seem to be present, for no cause is appealed to from without the agent, even though each of the agent's states is temporally ordered, and the ground of the change between states exists at a preceding time. Does Kant then lack any basis for rejecting a position that allows for temporally preceding causal grounds, so long as they are all within the agent?

In fact, Kant objects to just such a position as failing to put the levers of agency under one's control. Kant contends that such a position articulates a merely "relative" spontaneity—a "comparative" sense of freedom that he deems a "wretched subterfuge" (CPrR 5:96). Kant dismisses the possibility that this comparative sense could substitute for the control of one's actions required by morality.

in the question about that freedom which must be put at the basis of all moral laws and the imputation appropriate to them, it does not matter whether the causality determined in accordance with a natural law is necessary through determining grounds lying within the subject or outside him, or in the first case whether these determining grounds are instinctive or thought by reason; if . . . these determining representations have the ground of their existence in time and indeed in the antecedent state. (CPrR 5:96)

Here Kant's objection to the merely comparative freedom of being determined by one's internal states rather than an external force is that these internal states *would still be in time*, and thus that the determining grounds of one's act would

with here. Kant ultimately argues that God's creative activity is compatible with the freedom of created subjects, so long as time is not transcendently real. For discussion and evaluation of Kant's argument see Brewer and Watkins (2012), Hogan (2014), Kain (in press), McLear and Pereboom (in press).

lie in a previous time, and *thereby* outside of one's control. Note that here the argument does not rely on the way in which temporal causality traces back to an agent-external causal ground. Rather, the problem is temporal causality *itself*. Kant goes on to make this point quite clearly:

[purely internal temporal causes] are always *determining grounds* of the causality of a being insofar as its existence is determinable in time and therefore under the necessitating conditions of past time, which are thus, when the subject is to act, *no longer within his control* and which may therefore bring with them psychological freedom (if one wants to use this term for a merely internal chain of representations in the soul) but nevertheless natural necessity; and they therefore leave no *transcendental freedom*, which must be thought as independence from everything empirical and so from nature generally. (CPrR 5:96–97, original emphasis; see also *Ethik Vigilantius* 27:504 (1793/4))

Kant claims here that it is specifically the *temporality* of phenomenal causation that threatens the agent's control, and thus that even when the determining grounds of an agent's acts are all *internal* to the agent, the fact that they are in time is inconsistent with the agent's control with respect to them—the agent has no more freedom than a “turnspit” (*Bratenwender*, 5:97; see also *Metaphysik L₁*, 28:268–69).

But *why* is the temporal structure of causation itself a threat to the agent's control over her actions?²⁷ After all, on the above model the agent is, in every action, the ultimate causal source of that action, even though this causality is, as it were, “spread out” over the temporal history of her activity. Moreover, on this model there is no clear external threat to her agency in the manner that God's creative causal act of determining her nature and existence is an external threat.

One possible answer is that Kant is pointing in these texts to a second issue, distinct from that of agent-external determining grounds. Because of the nature of time, an agent acting at time t_n cannot causally affect the grounds g of that action, which exist at some prior time t_{n-1} , because (finite) agents cannot causally affect the past.²⁸ Hence, in the above passage Kant may be arguing that, putting

27. Sellars (1971: 20–21) interprets Kant as primarily worried about passivity with respect to the past and somewhat obscurely counters that “[t]he past is not something with respect to which we are passive” (1971: 21). Sellars seems to think that Kant conflates to some degree “foreign cause” with “temporally preceding cause.” If my interpretation is correct there is indeed a link between the notions of temporal cause and foreign/alien cause, but the link is genuine and not due to a conflation.

28. For a discussion of the problem of causal determination in terms of our inability to casually affect the past, and thus the scope of our causal reach see Sartorio (2013; 2016). Related to this issue, I think it is a mistake to read Kant's incompatibilism as based in something like Van Inwagen's “consequence argument”, though there is clearly some affinity between Kant's position and that sketched by the argument—see Van Inwagen (1983: 56); cf. Watkins (2005: 336–39) for discussion.

aside the issue of God's creative activity, so long as time is transcendently real one would lack freedom because one's acts would be due to grounds that are not within one's causal reach or influence at the time in which one acts.

However, one might worry about this being a compelling objection to a broadly Leibnizian (or Spinozistic) position, since at no time is the subject caused to act by a ground that does not lie within the subject. While it is clear that finite agents have no causal influence on the past, Kant does not give us further explanation of why this would undermine the central issue of the agent's control if all the relevant grounds of action are within the agent. Hence, it is not clear that emphasis on a lack of causal influence on the past provides a non-question-begging means of countering the position that freedom is compatible with determination by temporally preceding grounds.

I think we should thus consider an alternative reason for rejecting the compatibility of temporal causation with control, though this will require a brief excursus on the issue of Kant's conception of control over one's acts. Recall from Section 2 that Kant thinks of an action (*Handlung*) in its most general sense as belonging to the power or capacity of a substance to cause a particular accident (*Metaphysik Pölitz* 28:564–65). He then conceives of each faculty, whether receptive or spontaneous, as acting in its own specific way by producing specific characteristic effects, and thus operating according to specific and unique casual laws. The spontaneous acts of the intellect follow laws of "combination." Such acts are a form of "self-activity" (*Selbsttätigkeit*) and can never be given through the intuition of an object.

We saw in Section 2 that Kant's conception of a substantial act is very broad, encompassing any actualization of a substance's causal powers through an exertion of force to bring about an effect (a property or "accident") in itself or another being. While a substance's acts thus explain all of its accidents, Kant nevertheless wants to maintain the intuitive distinction between what happens to a substance versus what the substance does. He needs such a distinction, moreover, to distinguish between the two stems of our cognitive power—"passive" receptivity and "active" spontaneity.²⁹ He accomplishes this distinction by analyzing the relevant

29. Kant's active/passive distinction between sensibility and intellect can thus be somewhat misleading in this context, since he thinks that *any* determination in an individual is going to be the result of activity within that individual, even in the "passive" case of being affected by something else. This is clear from his discussion of substantial interaction in the *Metaphysics Mrongovius* lectures—see 29:823 and *passim*. In those lectures, Kant is reported as saying that "substance suffers (*passive*) whose accidents inhere through another power." He then asks, "How is this passion possible, since it was said earlier that it [i.e., the passive/suffering substance] is active insofar as its accidents inhere?" (*Metaphysik Mrongovius*, 29:823). Kant then asks, "What then is genuine passivity? The acting substance (*substantia agens*) determines the power of the substance being acted upon (*substantiae patientis*) in order to produce this accident, therefore all passivity (*passio*) is nothing more than the determination of the power of

cognitive faculty or power in terms of whether or not its exercise requires appeal to a “determining ground” whose nature or existence is distinct from that of the power in question, or whether the determining ground is itself part of the power’s exercise.^{30,31} Concerning acts that a subject does of itself, there may be further questions regarding whether those actions are voluntary or involuntary, which would then require understanding whether and how the will or power of choice is determined. But given this distinction between what happens to a subject and what the subject does we can speak (clearly if perhaps somewhat awkwardly) of “passive” or “receptive” acts of a subject in contrast to its “active” or “spontaneous” acts. For example, the sensory capacities individuated in characterizing our form of receptivity (e.g., vision, touch, etc.) all must appeal to determining grounds distinct from any exercise of those capacities.

In contrast, in the “active” acts of our spontaneous intellectual capacities, the existence and nature of the determining ground of the act is not itself independent of any exercise of the intellectual capacity. For example, in an inference, the determining ground of the inference, insofar as it is a rational act, must be the recognition of a logical relation between the premises, which is to say that the nature of the determining ground of the inference—that is, its status *as* a determining ground—itself makes reference to the very capacity for inference whose actualization is in question.³²

On this way of understanding Kant it should be clear why he would typically construe spontaneous mental activity as an action—in the sense of something an

the suffering substance by an outer power” (*Metaphysik Mrongovius*, 29:823). In other words, all change in a substance is at least partially due to the activity of that very substance, but some change is also due to the activity of some other individual acting on that substance. For further relevant discussion of the lectures see Thorpe (2011: 75–78), Stang (2019: 92–93). For contextualization in the Wolf-Baumgarten metaphysics of the day see Watkins (2005: 74–78 and §1, *passim*).

30. For related readings see also Engstrom (2006: 14–19), Smit (2009: 241–43).

31. What is a “determining ground”? It is best to understand this notion within its broader German philosophical context. According to Christian Wolff a determining ground is “that through which one can understand why something is” (*Deutsche Metaphysik*, §29; see also Baumgarten, *Metaphysica* §14). Kant rejects this as circular (NE 1:393) but maintains the general notion of that through which one understands or explains the existence of an object or its possession of some property (i.e., its being determined by some predicate). Note that a ground can be a complete (i.e., sufficient) ground for its consequent or merely a partial (i.e., insufficient) ground for its consequent. I take no stand here as to whether, in the exercise of a spontaneous capacity, the intellectual ground of such exercise is always complete, or whether the complete ground may include other partial grounds that are independent of the capacity’s exercise. For discussion of various elements of Kant’s conception of ground see Hogan (2009), Stang (2016a: ch. 3.2; 2019), Stratmann (2018), Watkins (2019b).

32. Another way to put this is that the very same formal cause is appealed to both in the individuation of the capacity and in the individuation of the determining grounds of its actualization or exercise. I say more about formal causation in Section 4.3 below.

agent *does*—since the intellect operates in such activity according to its own laws, and the determining grounds for its actions are themselves inextricably related to the nature of any exercise of the intellectual capacities themselves.

With this characterization of the nature of the receptive and spontaneous powers of the rational subject in hand, we can get a clearer sense of Kant’s notion of “control” and its incompatibility with temporally structured causation.

As I understand Kant’s notion of “control” (*Gewalt*), a subject is in control of their acts when the determining grounds of those acts are not independent of any exercise of the capacities for so acting. This means that an agent cannot be in control of any act that is itself the outcome of a determining ground whose existence and nature is itself independent of any exercise of the relevant capacity. Hence, all receptive acts are outside the agent’s control, which constitutes the sense in which they are passive “happenings” to the agent and the way in which something is “given” to it.

We can now see why Kant would reject the compatibility of the controlled exercise of a rational, spontaneous capacity with any temporal causal ground of its exercise or actualization. It is essential to the nature of a temporally structured determining ground that its position in time play a decisive role in the function of the ground to determine its particular effect. Hence the schematized law relating ground to consequence (as cause to effect) is one that makes essential reference to a temporal relation, and thus the position of each relatum in time. But then, even if the determining ground lies entirely within the agent, the sheer fact that it occurs in time means that the nature of that ground is, at least in part, independent of any exercise of the subject’s intellectual capacity or faculty, which itself makes no such reference to temporality. Thus, to construe a subject’s acts as brought about via a temporal causal ground is to construe those acts as, by definition, brought about via “alien influences” (*fremden Einflüssen*; G 4:448), that is, through grounds and laws external to one’s spontaneous capacities.³³ Moreover, since we distinguished the active and passive powers (i.e., receptivity and spontaneity) above in terms of whether the causal ground for their exercise ultimately depends on the relevant faculties, if an act is the result of a determining ground in time, then that act is due to an exercise of the subject’s passive *receptive* power.

Kant’s position thus stands as a genuine counter to the Leibnizian one rather

33. Here I anticipate issues that I discuss in further detail in Sections 4.2–4.3. For an alternative account to the one I provide here see Jordan (1969); cf. Hasker (1973), Allison (1996c: 98–104), Wolt (2018). An advantage of my approach is that it provides a clear explanation as to why Kant would find compatibilism (understood as the compatibility of natural determinism with freedom of the agent) such a non-starter. On my reading, it would be conceptually incoherent to endorse the compatibilism of rational action with temporal determinism given his conception of the nature of the intellectual faculty. For discussion of the issue of compatibilism see Ameriks (2000a: 19ff and 31ff; 2003b). I say more about the sense in which Kant must reject compatibilism below.

than simply begging the question against it. For Kant articulates a seemingly inescapable outcome of Leibniz's conception of the temporal structure of a substance's acts—viz. that since the determining grounds of the substance's acts are always temporal, the acts thereby necessitated will be *receptive* rather than *spontaneous*.³⁴ Put another way, this counter explains why Kant thinks the Leibnizian cannot recognize an *absolute* moment of spontaneity, but at best only the *relative* spontaneity of temporally preceding determining grounds that lie wholly within the subject but which are independent in nature and thus alienated from any exercise of its rational capacities.³⁵

Moreover, given this conception of control we can see why Kant would regard causation by an external source or entity (e.g., God) over which one has no causal influence or reach as threatening to control. Both the case of agent-external determination and the case of agent-internal temporal determination are instances of determining grounds that exemplify an independence incompatible with the actualization of an agent's spontaneous rather than its receptive capacities.

We can also now more fully appreciate why Kant defines transcendental freedom in the manner that he does. Recall that transcendental freedom consists in the (i) power to begin a causal series from oneself; (ii) without being determined to do so by any temporally structured causal ground. How should we understand the relationship between Kant's conception of control and his conception of transcendental freedom? One has control over one's acts just in case one is transcendently free, but the explanation of why Kant conceives transcendental freedom in the manner he does stems from the conception of control I have articulated. When one has control one exercises a capacity determined by a ground that depends on the very capacity so exercised. Hence one must be the causally ultimate source of the act. And for the reasons elucidated above the control one has in exercising such a capacity is incompatible with the ultimate determining ground of that exercise being in time. Moreover, the centrality of the notion of control for Kant's overall conception of freedom also explains why he

34. This seems true whether time is ultimately real for Leibniz or merely the well-founded phenomenon of some underlying order of relations between a substance's states. For even if time is merely a phenomenon, as well-founded it must be the expression of an underlying order whose relations are relevantly similar, hence insofar as the temporal nature of a determining ground indicates the independence of that ground from the capacity whose actualization it determines, so too must the non-temporal counterpart be independent of what it determines and thus, in Kant's terms, an alien influence. It is slightly less clear whether a conception of determining grounds according to Leibniz's later (roughly post mid-1690s) appraisal of time as merely "ideal" is susceptible to the same objection. For relevant discussion see McRae (1976: ch. 3), Hartz and Cover (1988), Adams (1994: ch. 9), Jauernig (2010), Jorati (2017: chs. 1–2, 6).

35. In this way Kant has a basis to object to even the specific kind of "rational spontaneity" Leibniz contends is characteristic of rational action, as opposed to the merely "metaphysical" and "agential" forms of spontaneity characteristic of non-rational acts. For discussion of these different kinds of spontaneity see Jorati (2017: ch. 2).

moves in the various texts I quoted above from conditions that undermine control to a lack of (transcendental) freedom.

3.3. *Leeway & "Ought"*

Thus far we've seen Kant's conception of freedom to be essentially causal, and concerned with the ultimate source of the causality of action, as well as the kind of control that the agent thereby exerts over those actions. However, one might presume from Kant's discussion that transcendental freedom also essentially involves the ability to do otherwise, that is, it essentially involves, in addition to the source condition outlined above, a *leeway* condition. Certainly, such a leeway condition is closely linked to Kant's conception of the power of choice (*Willkür*). One feature of Kant's definition of this power of free choice is that it is "a power to do or to refrain from doing as one pleases" (MM 6:213). Nevertheless, the concept of transcendental freedom does not entail leeway. A passage in the *Religion*, from which I've already partially quoted above, indicates this.

There is no difficulty in reconciling the concept of *freedom* with the idea of God as a necessary being, for freedom does not consist in the contingency of an action (in its not being determined through any ground at all) i.e. not in indeterminism ([the thesis] that God must be equally capable of doing good or evil, if his action is to be called free) but in absolute spontaneity. The latter is at risk only with predeterminism, where the determining ground of an action lies *in antecedent time*, so that the action is no longer in *my* power but in the hands of nature, which determines me irresistibly; since in God no temporal sequence is thinkable, this difficulty has no place. (Rel 6:50n)

As this passage shows, Kant does not conceive of the sort of transcendental freedom constituted by absolutely spontaneous action as requiring leeway in the performance of those actions. God's acts, including his acts of will, cannot be other than as they are, but this does not mean, in Kant's view, that God's actions are unfree. Despite lacking leeway in how he acts, God is absolutely spontaneous in the production of his actions, they are always within his control, and he is thus transcendently free. Hence, transcendental freedom in general does not essentially involve leeway or the ability to do otherwise. Kant is therefore a *source* rather than a *leeway* incompatibilist: the key notion of (transcendental) freedom is not the ability to do otherwise, but to be the undetermined causal source of one's actions, insofar as those actions are under one's control.³⁶

36. See Pereboom (2006a) for an important and influential statement of this kind of approach, though I do not claim that Pereboom would endorse all that I say here; cf. Kain (in press) for an approach that contrasts divine freedom with human freedom.

For Kant there is nevertheless an important connection between transcendental freedom and leeway, for it is precisely the fact that we can control our actions that makes possible the ability to do otherwise. Moreover, leeway of action is a necessary condition for the applicability of “ought” principles to us, and for us rational laws are presented via a system of “ought” principles. Kant says of the principles determining the will that,

The representation of an objective principle in so far as it is necessitating for a will is called a command (of reason), and the formula of the command is called an **imperative**. All imperatives are expressed by an *ought*, and by this indicate the relation of an objective law of reason to a will that according to its subjective constitution is not necessarily determined by it (a necessitation). (G 4:413; see also A547/B575)

Since the will (and ultimately the power of choice or *Willkür*) of a finite rational subject may or may not be determined by reason, the finite subject experiences the intellectual laws that might determine it as imperatives, and the application of such a law as a kind of necessitation. Since God is always perfectly determined by his rational intellect, and thus lacks any leeway in his actions, imperatives are correspondingly lacking.

no imperatives hold for the *divine* will and in general for a *holy* will: the “ought” is out of place here, because volition is of itself necessarily in accord with the law. (G 4:414)

One thing that is important to note in Kant’s conception of the conditions under which imperatives apply to rational subjects is that the relationship between the applicability of an “ought” claim and leeway need not require that choice always be *directly* operative in the occurrence (or not) of a specific act. For example, there are norms governing how one may adopt a particular doxastic attitude—what Kant calls holding-for-true or assent (*Fürwahrhalten*)—even though Kant denies that we can directly choose, irrespective of the epistemic grounds, whether or not to opine, believe, or know the truth of a judgment (JL 9:74; see also *Logik Blomberg* 24:158). Since leeway is a condition of the applicability of a norm, we possess leeway with respect to such attitudes not in the sense that we directly choose them, but rather in the sense that, if we are rational, then it is possible that such attitudes can be causally determined by rational rather than sensible causes.³⁷ One can thus “make up one’s mind” based on rational grounds rather than merely how one is sensibly determined. If it weren’t possible for the subject to adopt an attitude of assent based on rational causes then no

37. Thanks to an anonymous referee for encouraging clarity on this point. I discuss the issue of choice and intellectual acts further in Section 4.1.

ought would be applicable, and the rational norms would have “no significance whatever” (A547/B575).

Kant also endorses an “ought implies can” principle; e.g., “from the practical point of view this idea [of a prototype of humanity pleasing to God] has complete reality within itself, for it resides in our morally-legislative reason. We *ought* to conform to it, and therefore we must *be able* to” (Rel 6:62). If normative principles apply to us, it must be that we are, at least in principle, able to act in accord with them.³⁸ Hence, if one can never do otherwise because one is entirely (naturally) causally determined in one’s acts, it would be false that one ever *ought* to do otherwise. In this way the thoroughgoing determinism of the natural world threatens the applicability of “ought” judgments, understood as imperatives. If the thoroughgoing mechanism of nature entails that it is never true that one can do otherwise, there would be no point of a system of “oughts”, moral or otherwise. Hence, in order that “ought” statements for a subject be true at all, they must be true not only when the subject does x, but also when she *fails* to do x.³⁹ So, by Kant’s argument, “ought” principles cannot hold or be true if a subject could not have acted otherwise than she had. In order to have acted otherwise, however, it is not enough that the course of events leading up to and including her action could have turned out differently, in the sense that chance events might have occurred differently. Instead, an agent must have a capacity for action that is not subject to causal determination by factors beyond her control.

In sum, Kant construes free action as possible only if an agent has the requisite kind of control over her acts characteristic of the operation of her spontaneous intellectual capacities. Further, in finite rational agents, the applicability of norms as imperatives (i.e., as a system of “ought” claims) requires not only that the agent’s acts constitute or are otherwise a part of a causal sequence originating in the controlled exercise of spontaneous capacities, but also that she possesses leeway in her acts, with this latter fact explained by the agent’s being the ultimate causal source of her action.⁴⁰ Given these two conditions—viz. that an agent be transcendently free in that she is the appropriate *source* of her action and that she have the *leeway* in so acting necessary for being subject to normative principles formulated as “ought” claims—Kant’s conception of the temporal structure of phenomenal causality threatens the possibility of the truth of one or both of these claims.

The view thus far articulated explicates the incompatibility between Kant’s conception of a controlled use of one’s faculties and his conception of the mechan-

38. For detailed discussion of Kant’s conception of the “ought implies can” principle, see Timmermann (2003), Stern (2004), Kohl (2015a).

39. See Pereboom (2006a: 560 and note 40).

40. Thus Kant endorses two different requirements on the concept of free action, and sees the possibility of the second requirement as following from the first; see also Wood (1984: 77, 79–80), Kosch (2006: 17), Pereboom (2006a).

ical temporal causality of phenomenal nature. However, a number of questions remain. For example, to what extent are the activities of the intellect really free, especially in their theoretical use and what is the relation between such activities and will or choice? What reason is there for thinking that human beings instantiate the kind of control required for transcendental freedom? Even more generally, why think that such a conception of control is even possible for beings like us? In the next section I look at several different arguments concerning the transcendental freedom of the intellect's activity that employ the notion of cognitive control outlined here, and which attempt to provide answers, or at least the outline thereof, to such questions.

4. Elucidating the Freedom of Intellectual Activity

In this section I discuss four distinct though related arguments for the transcendental freedom of the intellect. Together they help to elucidate the nature of this free activity and serve as signposts outlining some of the major contours of Kant's theory of rational agency. I discuss objections to and criticisms of the arguments as they arise. In Section 5 I then address two more general objections to the position that the intellect must be, in both its theoretical and practical use, transcendently free.

4.1. The Imputability Argument

By far the most discussed argument as to why Kant is committed to some form or other of transcendental freedom in regard to even theoretical uses of the intellect is what I will call the "imputability argument."⁴¹ According to this argument, imputation of acts to a being such that they may be held responsible for those acts, and may thus be appropriate objects of praise, blame, and other "reactive" attitudes, requires that such beings be transcendently free.

We see this in Kant's distinction between ascription (*zuschreiben*) and imputation (*zurechnen*) in the following lecture transcription.

We can ascribe a thing [*einem etwas zuschreiben*] to someone, yet not impute [*zurechnen*] it to him; the actions, for example, of a madman or drunkard can be ascribed, though not imputed to them. In imputation the action must spring from freedom. The drunkard cannot, indeed, be held accountable for his actions, but he certainly can, when sober, for the drunkenness itself. (*Moralphilosophie* Collins 27:288 (1774/5); see also *Praktische Philosophie* Powalski 27:152–53 (1782/3), *Metaphysik* K₂ 28:739 (1790/91))

41. For recent extensive discussion of the issue of imputability and its connection to Kant's conception of doxastic freedom see Kohl (2015b); cf. Smit (2019).

In the lectures Kant uses the example of a drunkard whose actions while inebriated are not strictly imputable to him but merely ascribable (*Moralphilosophie Collins* 27:288). What *is* imputable to the drunkard is the decision to become (highly) intoxicated in the first place, for it is this action that Kant construes as under the drunkard's control, and correspondingly it is this action for which the drunkard can be held responsible, and to which his other actions while drunk must all be related. In the later *Vigilantius* ethics lectures Kant makes this point explicitly.

Auctor is an originator of action. Originator means that in regard to its determining grounds the action can, in its first beginnings, be derived from him. Hence he is regarded as the effectual first cause. . . . If it was a dizzy spell, then the cause was merely physical and a matter of natural necessity; it rested on no originative cause in the agent. If he was drunk, however, it was his doing to have gotten so; he knew the power of drink, and could have envisaged the possibility of evil consequences; he was thus the effectual cause, and it all began with him. (*Moralphilosophie Vigilantius* 27:559 (1793))

To the extent that the drunk was the “author” of his act—that is, made a free choice in deciding to get drunk—then the acts that are, strictly speaking, merely ascribable to him while thoroughly intoxicated are going to be things for which he is responsible because of the origin of those actions in the imputable decision to get drunk. We can thus say that the drunken actions are indirectly imputable to the drunk in virtue of this original act, but his blameworthiness for these acts is going to be directly dependent on his responsibility and blameworthiness for this original act.

At root here in the distinction between ascription and imputation is the notion of control. For Kant, ascription requires merely that one be the locus of some behavior or event, the determining grounds for which may be the laws of nature or something else entirely, as with a coerced or manipulated person. In such cases, we do not treat the actions of the coerced or manipulated (e.g., “brainwashed”) as under their control, and thus as appropriate targets of reactive attitudes or rational criticism. Similarly, non-rational animals may have acts ascribed to them but not imputed, for they are caught up in the deterministic mechanism of nature, which places the ultimate causal grounds of an act outside the agent's control.

As we saw in the previous section, Kant considers control a necessary condition of free action so it's unsurprising that he indicates its relation to the conditions of imputation. The epigraph to this paper records Kant as saying that

[if the human being] were led merely by natural laws, it would be impossible to impute to him any action, since the ground of action then would never lie in his control, but rather would be determined in the previous time. (*Metaphysik Vigilantius* (K₃), 29:1020 (1794/95))

Since we've seen that Kant's mature critical view considers control to be incompatible with one's acts being the result of temporally conditioned grounds or causes, it is clear that Kant assumes such control to be possible, if at all, only at the noumenal (i.e., non-temporal) level. One can therefore summarize the imputability argument as follows:

1. (Non-proleptic) praise/criticism of an agent for their actions is only rational if the agent is *responsible* for their actions.⁴²
2. Responsibility for one's acts requires that those acts be *imputable* to the agent rather than merely *ascribable*.
3. Imputation of action requires that an agent have the requisite *control* over her actions.
4. Control of action necessary for imputability entails the transcendental freedom of that action.
5. ∴ Non-proleptic imputation and related praise/criticism of an action requires that the agent performing the action be transcendently free.

There are at least two related worries concerning the above argument. The first is that considerations of imputability apply only to *voluntary* actions, and it is not at all obvious that intellectual acts, especially acts of synthesis, are voluntary. For example, it is plausible that Kant does not construe all acts of synthesis or concept formation as under the direct control of the will (*Wille*) or the power of choice (*Willkür*).⁴³ This worry thus affects the plausibility of premise (3), for one might object that the notion of "control" at work here requires the possession and relevant exercise of a will. If that's correct then only voluntary acts can satisfy Kant's conditions for imputability. The second worry concerns whether the imputability argument is suitable for showing a connection between imputability and specifically *theoretical* (as opposed to practical) exercises of intellectual freedom. This worry concerns the *scope* of the conclusion—can it apply to theoretical acts as well as practical ones? I'll take these in turn.

42. By "non-proleptic" praise/criticism, I mean such praise or criticism that is not merely forward looking, with the aim of changing behavior. One praises or scolds a dog, not because it is blameworthy, but because one wishes to shape its future behavior. The applicability of reactive attitudes as Kant considers them are *both* forward looking, towards future behavior, and backward looking, towards the original decisions that are the source of the behavior. For discussion see Williams (1995: ch. 6), Pereboom (2014).

43. For example, if imputation is closely tied to the conditions for praise and blame, it might be thought implausible to regard, e.g., figurative synthesis as something for which one could be rationally praised or blamed. However, something like this account is defended in contemporary philosophy. For example, Siegel (2017b) argues that various sorts of bias can affect the rational standing of one's experiences, and thus that one can be rationally criticized and appropriately be the target of reactive attitudes for experiences that one has. As we'll see below, Kant's stance regarding such a position is not entirely clear.

First, does imputation require that the relevant act always be under *voluntary* control, and thus a product of the exercise of the agent's will, or can even non-voluntary acts be under one's control?⁴⁴ More specifically, the question is whether imputation requires volition or an act of will at all.⁴⁵

In my view there is no specific text of Kant's that definitively settles this issue one way or another. Instead, we need to work out an account that best fits with his explicitly expressed positions. One case of an intellectual act that Kant does explicitly discuss in relation to the exercise of the will is that of doxastic attitude formation or assent (i.e., "holding-for-true" or "*Fürwahrhalten*").⁴⁶ Though it is somewhat controversial, Kant appears to straightforwardly deny that doxastic attitude formation is under one's immediate voluntary control. For example, in the *Jäsche Logik* and related lecture notes Kant holds that the "will does not have any influence immediately on holding-to-be-true" (9:74; see also BL 24:156).

Kant's conception of the "immediate" connection or influence of the will on assent is not thoroughly spelled out. As he explains it, the question is whether a phrase like "gladly believe"

[would] seem to indicate that there is something elective [*etwas Willkürliches*] in our judging, in that we hold something to be true because we want to hold it to be true. (JL 9:73)

Kant immediately goes on to deny this claim,

If the will had an immediate influence on our conviction concerning what we wish, we would constantly form for ourselves chimeras of a happy condition, and always hold them to be true, too. But the will cannot struggle against convincing proofs of truths that are contrary to its wishes and inclinations. Insofar as the will either impels the understanding toward inquiry into a truth or holds it back therefrom, however, one

44. Two somewhat ancillary points are nevertheless worth keeping in mind here. First, we shouldn't equate volition with the possession of leeway—i.e., the ability to do otherwise. For as we saw in the previous section, the two can come apart. God cannot do otherwise than he does, but God's acts are both voluntary (i.e., the outcome of God's will) and imputed to him. Second, even for non-voluntary acts, i.e., acts not determined by an agent's will or power of choice, it still makes sense to ask whether such acts are transcendentally free. In particular, it makes sense to ask whether the act constitutes or is a result of a causal chain initiated by the agent and through or over which the agent exercises control. For a related attempt to secure a causal role for the agent in intellectual acts see Pereboom's (1995: 4) response to Kitcher's (1990: 122) position that all acts of synthesis are "subpersonal" in Daniel Dennett's (1969: 93–96) sense.

45. For a contemporary example of an account of imputation without volition—instead, the key notion is one of "reactivity" to reasons—see McHugh (2012).

46. The will's relation to assent has been much discussed; for examination of some of the relevant issues see Alston (1988), Stevenson (2004), Smith (2005), Cohen (2013; 2014), Kohl (2015b), McCormick (2017), Vance Buroker (2017).

must grant it an influence on the use of the understanding, and hence mediately on conviction itself, since this depends so much upon the use of the understanding. (JL 9:74; see also BL 24:157–59)

Kant denies the possibility of assent “at will,” which he construes as the immediate byproduct of choice, but allows that one can will oneself into picking up or dropping the investigation of the truth of some judgment, as well as the means by which such an investigation or inquiry is accomplished.⁴⁷ I take Kant’s denial here to, at the least, entail the denial that an act of will or exercise of the power of choice could be, by itself, the sufficient cause of one’s adopting, rejecting, or otherwise revising one’s assent to a judgment. In this sense Kant rejects what is now called “direct doxastic voluntarism.”

Kant’s rejection of direct doxastic voluntarism is important for answering the first objection in the following respect. Kant considers reflection on the grounds by which one assents to be, at least on some occasions if not generally, a “duty from which no one can escape” (Amphiboly, A263/B139; see also JL 9:76; BL 24:161). He also thinks that accepting something while deliberately avoiding further reflection is a kind of “punishable prejudice” (*sträfliches Vorurtheil*). It thus is plausible that Kant holds that assent to a judgment is an imputable act for which a subject can be held responsible, and for which she can reasonably be praised or blamed. Since Kant holds that assent is not, at least typically, something that is the immediate outcome of will or choice, Kant must not construe imputability as relying directly on one’s volition or power of choice.

What is much less clear is whether Kant construes imputability to require what he calls an “indirect” or “mediate” exercise of the will. If he doesn’t, then a lack of connection to volition is no problem for the general imputability of intellectual acts. But even if Kant does require that imputable acts have at least a mediate connection to the will, I think he can accommodate this via his conception of the centrality of attention for all rational activity.

When I introduced the notion of attention, in Section 2, as a kind of intellectual act, I argued that Kant understands attention as an act in which a subject can exert control over their consciousness of the various determinations of an object. This control over one’s representations comes from the subject in that it is independent of representational stimulus, cause, or occasion. Kant’s somewhat crass example is that of a suitor who “could make a good marriage if only he could overlook a wart on his beloved’s face, or a gap between her teeth” (An 7:131–32). The idea being, clearly, that the suitor would be happier if he were to exert attentional control over aspects of his representations, ignoring some elements while focusing on others. It is precisely this ability to exercise one’s will in controlling one’s

47. For this point, as well as discussion of the above passages, see especially Cohen (2013: 35–36).

attention that is lacking in non-rational animals. The attentional capacities of non-rational animals are determined in them by nature. The threats and promises that populate such an animal's environment, along with its natural needs and inclinations, push and pull on its attentional resources, resulting in pathologically determined inclinations to act in various ways. This is why Kant construes the animal power of choice (*arbitrium brutum*) as "pathologically necessitated", in contrast to the free (*liberum*) power manifested by rational beings (A533–34/B561–62; see also *Metaphysik L₁* 28:255 (c. 1778–1781); *Metaphysik Mrongovius* 29:896 (1782/3); MM 6:213).⁴⁸

Hence, even if Kant regards free acts as being or involving the exercise of the will as ground of the power of choice, the central role of attention in acts of conceptualization, judgment, and reasoning allows the will a pivotal place in the determining grounds of a rational agent's actions, even those that are not under its direct voluntary control, such as assent.⁴⁹ For this reason, I think the imputability argument might plausibly be extended to cover not just overt intellectual acts such as holding-for-true, but also acts of judgment, conceptualization, and perhaps other forms of synthesis as well.

Moving now to the second objection mentioned above, the worry is that Kant's account of imputation derives entirely from his conception of *practical* or *moral* freedom, rather than from any feature of the intellect's theoretical exercise. Hence, regardless of how we settle the issue of volition with respect to intellectual acts, the worry is that there are no specifically theoretical or cognitive norms. Instead all norms are inherently practical or moral. Patricia Kitcher puts the point this way,

Intellectual accountability is thus an important aspect of [Kant's] views and of his legacy. Although this point is clearly correct, it does not imply that he thinks that there is a special sort of accountability involved in cognition. Rather, his view is that cognizers must use their cognitive faculties in accord with the moral law. (2011: 247)

it is not errors of theoretical reasoning per se that are imputable, but failures to develop reasoning skills and to take up controversial topics that can be faulted. (2014: 158)

48. For further defense of this claim about attention and its connection to rational activity see McLear (b; c); cf. Merritt and Valaris (2017), Merritt (2018: ch. 3).

49. It's worth emphasizing here that the central role accorded to attention does not require that Kant adopt any view stronger than an indirect doxastic voluntarism, according to which intellectual acts (e.g., judgment, assent) are not determined "at will" but rather affected by upstream willful acts (such as what we attend to and how). For discussion of Kant's endorsement of indirect voluntarism see Cohen (2013), Vance Buroker (2017). Thanks to Eric Watkins for urging clarity on this issue.

If Kitcher is correct then there is no reason to think that imputability in the case of rational criticism of a thinker amounts to anything more than the moral requirement that the agent develop their (rational) talents (Kitcher 2011: 248).

However, Kant's remarks in the Canon concerning assent indicate a more direct relation to laws governing imputability of assent than Kitcher's interpretation allows. For example, he says that,

I must never undertake **to have an opinion** without at least **knowing** something by means of which the in itself merely problematic judgment acquires a connection with truth which, although it is not complete, is nevertheless more than an arbitrary invention. Furthermore, the law of such a connection must be certain. For if in regard to this too I have nothing but opinion, then it is all only a game of imagination without the least relation to truth. In judging from pure reason, **to have an opinion** is not allowed at all. (Canon, A822/B850; original emphasis)

Kitcher might argue that the normative requirement in the first clause not to opine without knowing something relevant to the opinion's truth is simply an expression of moral norms governing the development of proper reasoning skills. However, the rest of the passage indicates that the source of the normative requirement is the structure of assent *as such*. One cannot even count as opining (a form of assent) unless one's opinion has a connection with the truth. Otherwise one merely plays a "game of the imagination." This is not merely a condition related to developing one's (rational) talents or skills. It is a requirement on assent.⁵⁰

If these replies are cogent then the imputability argument provides at least some reason to think that any rational agent that is correctly open to criticism with respect to intellectual acts such as those of conceptualizing, judging, inferring, or assenting, must, by Kant's lights, be transcendently free. Note that the argument does not show that *we* are agents whose acts are imputable. The best the argument shows is the truth of the conditional claim that *if* agents are to engage in or correctly be objects of non-proleptic criticism or praise *then* they must exhibit control over their actions, and thus be transcendently free. Moreover, the imputability argument does not provide a complete picture of Kant's reasons for thinking that intellectual acts must be free. To get a better sense of Kant's underlying view I examine below three further arguments for linking freedom

50. If there are constitutive laws governing rational assent, and more generally, governing all rational connections between mental states, then a further commitment of this interpretation of Kant is that it construes the rational commitment present in a "hypothetical imperative" as itself dependent on constitutive rational laws that manifest themselves to finite rational beings in the form of categorical imperatives. For argument that Kant must construe the relationship between hypothetical and categorical imperatives in this manner see Korsgaard (1986; 1997), Engstrom (1993). Thanks to Eric Watkins for urging clarity on this point.

and intellectual activity. The first concerns his conception of how reasoning works, which we'll see requires that the intellect be transcendently free in its activity.

4.2. *The Reasoning Argument*

Kant's "mark" (*Merkmal*) theory of concepts construes concept possession in terms of recognition of conceptual content as the "ground" (*Grund*) of one's cognition of a thing.⁵¹ As he puts it,

A mark is that in a thing [*Ding*] which makes up part of its cognition, or—what is the same—a partial representation so far as it is considered as ground of cognition [*Erkenntnisgrund*] of the whole representation. (*Jäsche Logik* Introduction §8; 9:58)

Representation of a ground *as such* requires the generation of or transition to a mental state by recognizing the content of one's present or previous state as the basis of the new state. Recognition of a ground of difference (or, relatedly, of similarity) is, for Kant, a recognition of a rational relationship between two (or more) things or properties. To take a simple example, one *rationally* distinguishes between two differently colored objects not merely in the sense that one's visual system is sensitive to differences in color but rather because one also recognizes that one object is *blue* and the other *red*. In so recognizing this basis for distinguishing the objects one takes the color of the first object as a ground of its being different from the second object, and thus a reason for judging in one way rather than another. That is, in differentiating one object from another by their color one recognizes that, for example, blue objects cannot simultaneously be red and vice versa. Similarly, in thinking that an object is red one also thereby recognizes that its redness is reason to think it is colored and extended in space.⁵²

It is crucial here that the subject not simply have states that flow one to the next in a manner that respects their content. It has to be the case that the subject appreciates these connections *as such*, and thus that her mental states are formed on the basis of such an appreciation of their content. Thus, the cognitive subject must play a causal role in the production of her states by recognizing that the content of one mental state connects with that of another and producing this other state on that basis.

⁵¹. For discussion of representational marks generally see Smit (2000), Grüne (2009). For a recent elaboration of this point see Kitcher (2011: 120).

⁵². One way this has been put in the secondary literature is in terms of the "inferential integration" of one's mental states. See Pereboom (1995: 17). A similar position is advocated by Kitcher (2011: 120ff). Such integration of one's mental states is also often considered a marker of the "doxastic" as opposed to the "sub-doxastic" (e.g., Stich 1978: 506–507) or the "personal" as opposed to the "subpersonal" (e.g., Bermúdez 2005: 31).

Conceptual representation thus requires that the conceptualizing subject be able to base one state (or act) on another in the above manner. Kant distinguishes this kind of mental transition (call it “reasoning”)⁵³ from other kinds of mental transition (e.g., association in its various forms) in virtue of whether the subject has so based the succession of her states in a manner that recognizes the content of the one state as the basis for another. Following a variety of others, let’s call this condition on rationally connecting one’s states the “taking” condition, which says that a necessary condition of rationally basing one mental state (or act) on another—as happens in the inferential transition from one representational state to another—is that the subject take her state to be formed on the basis of her other state(s).⁵⁴

If Kant does in fact endorse a taking condition on reasoning, and thus on conceptualization, then we can see how the variety of rational activities that he typically categorizes in terms of “synthesis” or “combination” (e.g., conceiving, judging, inferring, and assenting) are all going to be related via the central conception of basing one contentful state on another. We can then formulate an argument that such takings must be transcendently free as follows.

1. Conceptual representation constitutively depends on the exercise of one’s capacity for reasoning—that is, the capacity for a particular kind of transition between one mental state and another.
2. Reasoning requires “taking” one state as the basis for the next.
3. Taking is an act under the subject’s control.
4. If the transition from one state to another is merely part of the “mechanism of nature”—that is, it is wholly determined by temporally preceding causes—then it is not under the subject’s control.
5. ∴ Mental transitions due to the mechanism of nature are incompatible with the subject’s engaging in reasoning—that is, reasoning requires transcendental freedom.

53. I intend “reasoning” to be used more broadly than *sylogistic* reasoning or inference, as it is understood by Kant. “Reasoning”, in the sense at issue here, concerns the activity characteristic of conceptual representation, judgment, and syllogistic inference. All three types of mental activity are reasoning in the sense of requiring a consciousness of one mental state as based on others.

54. The claim that there may be such a “taking” condition on rational transitions between one’s mental states or acts is sometimes attributed to Frege—e.g., Boghossian (2014: 4). That there is such a taking condition on rational transitions, particularly those characterizing inference or reasoning more generally, is controversial. For discussion and criticism see Broome (2013; 2014), Wright (2014), McHugh and Way (2016), Siegel (2017a). Here I shall say relatively little to defend the taking condition as such. For further defense and discussion of the taking condition as it relates to Kant’s views concerning apperception and reflection see McLear (b). For contemporary defense of an account of the basing relation that is broadly similar to the account I have in mind here see Neta (2019).

That there is a link in Kant's view between "takings" and transcendental freedom has been, at least to some degree, recognized in the secondary literature.⁵⁵ What is less clear is why the act of taking is linked to such freedom. I've argued for why Kant accepts premises (1) and (2). We've also seen, in Section 3.2, that the temporally conditioned causal structure of nature is incompatible with the subject's actions being under her control, and thus that Kant accepts (4).⁵⁶ The key then is to understand why he might accept (3). One reason might be that he accepts that we can hold people culpable for instances of poor reasoning, and thus that he relies on the imputability argument for justifying the acceptance of (3). But there are also considerations distinct from imputation, which stem from Kant's conception of a faculty, that support his endorsement of (3). I discuss these in the next subsection.

55. See, e.g., Paton (1947: 218), Pippin (1987: 468), Allison (1990: 37–39; 1996b: chs. 4, 7 and 9; 2006: 389), Willaschek (2006: 169–71); cf. Ameriks (1992; 2003a: ch. 6). Allison has perhaps the most extensively developed view of the spontaneity of the intellect—one that is in many ways congenial to the account I have developed here. However, there are some important differences. For example, Allison characterizes Kant's conception of freedom as essentially "Crusian", which is to say that freedom depends on the capacity to do or refrain, and thus on leeway (see Allison 2006). But I've argued that Kant's view is at root a *source* view of freedom, and hence of leeway being a derivative or downstream effect of the source view, being necessary for explaining the applicability of "ought" claims but not sufficient for freedom as such. Relatedly, Allison claims that "Kant regards the capacity to act on the basis of imperatives . . . as the defining characteristic of free agency" (1990: 36) and correspondingly as making "the virtual identification of rational agency with action on the basis of an ought" (1990: 38). This again, in my view, puts far too much emphasis on leeway. Perhaps more important though, while Allison recognizes that Kant often speaks of reason, or the intellect broadly construed, as a causal power (e.g., Allison 1990: 47–48), Allison nevertheless seems inclined to reject this reading. He instead appears to favor a view where reason (here specifically concerning incorporation of maxims, which involves a version of "taking" as discussed above) "has causality" only in the Pickwickian sense that it provides the guiding rule" (Allison 1990: 51; cf. Ameriks 1992: 215). In my view, Kant is committed to the causality of the intellect in a wholly literal sense, though there may be different kinds of causality in play—see Section 4.3 below for further discussion. Moreover, it is precisely because Kant is committed to a non-Pickwickian sense of the causality of the intellect that he must reject any view of the rational mind that construes its acts as fundamentally temporally structured and thus wholly subject to phenomenal causation. Finally, and related to this point, Allison denies that the freedom of the intellect plays any metaphysically explanatory role in Kant's theory—construing it instead as a wholly conceptual claim (see, e.g., Allison 1996b: 126, 142). In my view, while it is true that Kant conceives of the intellect as a free causal power, this in no way inhibits Kant from construing the freedom of the intellect as importantly explanatory, as the taking argument clearly indicates.

56. For an alternative account of Kant's argument for (4) see Wolt (2018).

4.3. The Constitutive Laws Argument

As discussed in Sections 2 and 3, Kant conceives of the various faculties and sub-faculties of the mind in terms of their causal activity. That is, for each faculty there is a characteristic set of effects for which it is responsible, and thus a characteristic set of acts that bring about these effects. Since causality is at issue in describing the relation between an act and its effect, the characterization of a faculty in terms of causal activity requires, for Kant, characterizing that faculty in terms of the constitutive causal laws governing that activity.

Because the laws governing the acts of a faculty are constitutive, the faculty, when not interfered with, always generates effects in accordance with those laws, for this is just what it is for the faculty to be what it is.⁵⁷ Kant is very clear about this point in the Transcendental Dialectic.

No force of nature can of itself depart from its own laws. Hence neither the understanding by itself (without the influence of another cause), nor the senses by themselves, can err; the first cannot, because while it acts merely according to its own laws, its effect (the judgment) must necessarily agree with these laws (Transcendental illusion, A294/B350; see also JL 9:53–54, *Wiener Logik* 24:824–25, R2142 16:250 (1776–81))⁵⁸

Kant construes error as the result of the deviation of a faculty's activity from its characteristic laws in a kind of "geometry of error", and which, since the issue is always one of causality, must be construed in terms of the causal interaction of different faculties upon one another.⁵⁹ As he puts it in the *Vienna Logic*:

57. See also Kant's discussion of the Holy Will in *Groundwork I*, 4:413–14, part of which was discussed above in Section 3.3. Of course, since we have a plurality of faculties, the effects of those faculties can cause "deviations" (*Abweichungen*) from their normal effects, and thus errors. See the account of error below.

58. Note that the appeal to "nature" (*Natur*) in this passage is nature in its "formal" rather than "material" sense, where this concerns "the first inner principle of all that belongs to the existence of a thing" rather than the "the sum total of all things, insofar as they can be *objects of our senses*" (MFNS 4:467). It is clear both from context, and the various logic lectures where Kant makes the same points about error, that his discussion concerns "essential laws" (e.g., *Wiener Logik* 24:824) of the sensible and intellectual faculties—i.e., laws articulating their nature in the formal sense. Moreover, Kant's discussion of a "force of nature" (*Kraft der Natur*) is meant to indicate the relation that a substance bears to the accident whose existence it causes; see Longuenesse (1998: 7–8), Dyck (2014: 200–201), Stang (2019: 92–93). So Kant's discussion in the text of the Dialectic as well as the lectures concerns the respect in which the actualization of a specific faculty in a substance causes the existence of a particular accident (in this case a possibly erroneous judgment). For discussion of the nature of error see Butts (1997: ch. 4), Grier (2001). Thanks to an anonymous referee for encouraging clarity on this point.

59. There are a variety of different ways one might characterize this influence. See Butts (1997), Tolley (2006: 385; 2008: 224–25), Rödl (2007: 117–20), Frierson (2014: chs. 3 & 6), Kohl (2015b: §3), Merritt (2018: ch. 2) for relevant discussion.

Error is neither in the understanding alone, then, nor in the senses alone; instead, it always lies in the influence of the senses on the understanding, when we do not distinguish well the influence of sensibility on the understanding. (24:825)

Errors in thinking, for Kant, arise as a product of the influence of sensibility on the intellect.⁶⁰ So, the laws of the intellect produce particular effects unless interfered with by some external or “alien” causal power.

What are the laws by which the intellect operates? At the most general level they must include the laws of logic. Kant says,

[Pure general logic] contains the absolutely necessary rules of thinking, without which no use of the understanding takes place. (On logic in general, A52/B76)

These logical laws must be operative for the occurrences of and transitions between the mental states of a subject to count as genuine cases of *thought* or thinking. This means that in cases in which those laws are not operative, such as when the understanding’s activity is brought about through an alien cause, the status of these deviations is not characterized by Kant as thinking.⁶¹ For example, in a note from a series discussing Meier’s conception of error in his *Auszug* Kant speaks of the “mixed effect” (*vermischte Wirkung*) generated by the influence of sensibility on the intellect, an effect that one would err in calling a “judgment of the understanding” (R2244, 16:284; see also JL 9:73).⁶²

Kant also indicates in other published work that one’s rational states are generated according to the laws constitutive of the rational faculty’s operation. For example, in the *Groundwork* he famously says that,

one cannot possibly think of a reason that would consciously receive direction from any other quarter with respect to its judgments, since the

60. There must be a caveat here since, on Kant’s view, some errors arise in a manner that is internal to the intellect itself, in the dialectical structure of reason’s search for systematicity by relating condition and conditioned. However, this view is still consistent with Kant’s general account since such dialectical errors arise via the relation of sensibility to reason’s search for an unconditioned ground. In this sense all dialectical error involves a “subreption” in which concepts or principles meant solely for use in application to appearances in experience are applied (whether knowingly or not) outside of experience. Error thus does not arise simply from the *generation* of the ideas themselves, but rather only in their “misuse” (A669/B697). See Grier (2001: chs. 2–4), Rohlf (2010), Dyck (2014: ch. 3.2).

61. For other defenses of this kind of “constitutive” approach to thought see Thompson (1981), Conant (1992), Tolley (2006; 2008); for a more “normative” approach see MacFarlane (2002), Longuenesse (2005: ch. 4), Leech (2015), Lu-Adler (2017), Pollok (2017), Nunez (2018).

62. See Meier (2016: §§36–65) and Kant’s reflexionen on Meier’s position consisting of R2242–R2274.

subject would then attribute the determination of his judgment not to his reason but to an impulse. (*Groundwork III*, 4:448; see also CF 7:27)

Given Kant's constitutive and causal conception of the activity of a faculty, it is understandable why he would say that reason cannot consciously "receive direction" (*eine Lenkung empfinde*) from something outside itself and still count as reasoning or judging, for *what it is* to reason, judge, or in general to think, is to generate specific kinds of mental effects in accordance with laws characteristic of one's intellectual faculty rather than in accordance with laws as determined by something else distinct from that faculty.⁶³

Hence, if the ultimate ground or cause of a mental state, or of the transition from one mental state to another, consisted in a causality originating in a power other than that of the intellect, the kind of "mentation" involved (if one can forgive the neologism) cannot be, for Kant, genuine *thought*. This means that, for thought to occur, the ultimate ground or cause of one's thoughts cannot be due solely to the "mechanism of nature," understood as the temporally structured causality of the natural phenomenal world. If one's mentation *were* due to natural causal laws, then it would be *ipso facto* due to "alien influences" (*fremde Einflüssen*), and thus due to a causality distinct from that of the intellect.

We now have the material for the following argument, which makes these commitments explicit.

1. The acts of a faculty are governed by its constitutive laws.
2. The laws governing the intellectual faculty (*der Verstand; Intellectus*) are not those constituting the mechanism of nature.
3. If the ultimate ground of a transition from one mental state to another is due to the mechanism of nature then it would not be a transition determined by intellectual laws.

63. See also Paton (1947: 218, 220–21), Johnson (2009: 100). Now, one might object that at G 4:448 Kant only speaks of what reason is *conscious* of with respect to judgment. Might it not be the case that he allows that reason could *unconsciously* receive direction from some non-rational source? In other words, is there some form of determinism that could be made compatible with Kant's claim? At least two points are relevant in reply. First, since reason's activity is inherently self-consciously accessible (this is Kant's *cogito* condition stated most clearly in the B-Deduction at B131–32), there can be no ground of judgment relevant to the judgment's rational assessment that could be in principle inaccessible to reason; see also Paton (1947: 220–21); cf. Jordan (1969: 51–53). Second, another way of stating my point in the main text is simply that if reason receives direction from another source, then consciously or not, it is not *reason* that is serving as the ultimate causal explanation of why the agent judges, infers, etc. It is thus part of the very concept of reason (or the intellect more broadly) that it play this ultimate role. Here I agree with Allison that there is an important conceptual connection for Kant between reason or the intellect and freedom, though I disagree with Allison that this means that the freedom of intellect does not play any explanatory function; see Allison (1996b: 126 and note 47).

4. ∴ If a mental transition is intellectual/rational—that is, if it is to count as *thinking*—it must be both lawful and independent of the mechanism of nature and thus transcendently free.

We can see now why, in the “taking argument” discussed previously, Kant must claim that taking (and thus reasoning) is something under the subject’s control. If taking were not under the subject’s control then an alien causality would be at work, and the resulting mental activity or state could not thereby be characterized as a case of thinking. The constitutive laws argument also helps us understand why the imputability of an act is so closely linked to the causal conditions of its production.

One might object here that intellectual acts cannot be transcendently free in the way necessary for satisfaction of the constitutive laws argument, because they cannot plausibly constitute or be part of a causal chain initiated wholly by the subject. Patricia Kitcher puts the objection this way,

[consider] the inference: “if *p*, then *q*,” “*p*,” therefore “*q*.” Insofar as a transcendently free cognizer sees his judgment “*q*” as a new beginning, he would not see it as dependent on his other representations, and hence would not through the exercise of his cognitive faculties come to understand his states as belonging to the unity of apperception. He would have no grounds for seeing his states as necessarily connected. (Kitcher 2011: 169, see also 240)

In reply it is important to see that this objection locates the initiation of a new causal chain in the wrong place.⁶⁴ Kitcher assumes that the acceptance of *q* stands at the beginning of a new causal sequence, and rightly worries that this would cut off the conclusion of an inference from its sufficient basis, invalidating the supposed rationality of the inference. But there is no need to interpret Kant’s claim in this way. It is not the acceptance of *q*, but a different mental act—the *taking* of *p* and *if p then q* as the basis for accepting *q*—that stands at the beginning of a new causal chain. This act is the act of relating premises to the acceptance of a new judgment, and if the ultimate causal ground of such an act is merely the mechanism of nature, or some other alien cause, then the result of that activity—viz. the transition to *q*—would not be an inference. It must be the causality of the intellect that is ultimately the ground of the inference to *q* and not merely that the

64. An alternative is to see Kant as arguing that rational necessitation is not causation at all. As Kohl (2015b: 318 note 44) puts it, “the idea that proposition A entails proposition B as a matter of *rational* necessity does not entail that a thinker who accepts A is *causally* necessitated to accept B.” However, if I am right as to how Kant construes the nature of the faculties, it is incorrect to construe rational necessitation as non-causal. It is simply causal in a different manner from that of the temporally structured causality of nature.

preceding state of physical or psychological reality was thus-and-so.⁶⁵

One might further object to the construal of a faculty's operations in terms of causality on the grounds that it is implausible to understand logical laws as *causal*. In reply, I think we have to explore different notions of causality as they work in Kant's system. The judgment that is the effect of an act of a faculty is, I think, supposed to be *efficiently* caused by the being (i.e., substance) whose faculty that is. As noted in Section 2, Kant construes faculties as the causal powers of substances to efficiently cause accidents, either in themselves or others. This is not a view that is unique to Kant, and goes back at least to Leibniz, who also holds that substances are the efficient causes of their states or accidents through an exercise of their causal powers.⁶⁶ But the fact that a particular judgment is efficiently caused by an act of judging does not rule out the possibility that Kant also thinks there are other kinds of causation necessarily involved in judgment. Indeed, in my view Kant also has to appeal to both formal and final causation. Let me briefly discuss final causation before turning to formal causation.

Final causation comes with what Kant calls the "interest" of a faculty:

To every faculty of the mind one can attribute an interest, that is, a principle that contains the condition under which alone its exercise is promoted. Reason, as the faculty of principles, determines the interest of all the powers of the mind but itself determines its own. The interest of its speculative use consists in the cognition of the object up to the highest a priori principles; that of its practical use consists in the determination of the will with respect to the final and complete end. (CPrR 5:119–20)

Part of the ultimate explanation of a rational connection of one's states concerns the fact that they are connected for a particular end determined by the fundamental interest of one's rational capacities in their speculative use—viz. achieving cognition, and ultimately what Kant calls "comprehension" (*Begreifen*), via a priori principles (JL 9:64–65; see also *Dohna-Wundlacken Logik* 24:730–31 (c. 1792); *Wiener Logik* 24:846 (1780); *Blomberg Logik* 24:132–33, 134–35, 136 (c. 1771)). So, that for the sake of which all rational transitions involving the speculative (or theoretical as opposed to practical) use of our rational faculties are made is the greater understanding or making sense of things ultimately perfectly manifested in comprehension.

That the effects of a rational faculty have such a final cause says nothing about the structure in terms of which a determinate effect—e.g., a cognition—is

65. Wood (1999: 175) argues in a manner that may be congenial to what I have suggested when he says that "Kant holds that we must think of ourselves as free in all our rational judgments in the sense that we must regard our judgments as acts we perform under norms." However, it isn't entirely clear whether Wood construes performance "under a norm" in terms of a rationally constitutive causal law. Insofar as he does, our interpretations are compatible.

66. For discussion of Leibniz's views on causality see Jorati (2015a; 2015b).

efficiently brought about by the substance that has that faculty. It doesn't, for example, say anything about the very different ways such an effect might be brought about by/in an intuitive intellect vs. a merely discursive one. But the appeal to the kind of structure necessary to so differentiate the two kinds of ways (viz. intuitive vs. discursive) of achieving the end of cognition looks to be just an appeal to the formal cause of the cognition. If this is correct, then Kant has to appeal to what looks like formal causation to give a full explanatory story of a determinate effect, like a judgment, being brought about according to intellectual laws (i.e., by rational causes). This would then mean that any rational act or transition in the mind has a formal cause, and that this would bottom out, at least for discursive minds, in Kant's account of the unity of apperception, which is the intellectual form of all discursive cognition.

If this is all correct then while the intellect is that causal power of a substance through which it efficiently causes particular kinds of effects in accordance with laws of combination, a full explanation of the nature of this causal power and its effects will appeal to more than just efficient causation. Rather, such an explanation will also appeal to final causation (i.e., the aims of the cognitive discursive mind) and formal causation (i.e., pure apperception as the form of all intellectual cognition).

Kant's position that the spontaneous exercise of one's intellect works according to constitutive causal laws distinct from those of nature also helps shed light on one of his more perplexing arguments against hard determinism (or "fatalism"). The argument appears in Kant's 1783 *Review* concerning Johann Heinrich Schulz's work on ethics. There Kant says that,

[The fatalist] has assumed in the depths of his soul that understanding is able to determine his judgment in accordance with objective grounds that are always valid and is not subject to the mechanism of merely subjectively determining causes, which could subsequently change; hence he always admits freedom to think, without which there is no reason. (8:14)

Kant's argument here is highly compressed. But one point he might be construed as making is that the proponent of fatalism undermines her own position because she can provide no guarantee that her judgments are based on reliable objective grounds rather than unreliable subjective ones, which as Kant points out "could subsequently change." On this reading Kant thus argues that the fatalist lacks any justification for her holding true any of her judgments, for such judgments are just the lucky result of the deterministic order of nature. One problem with this reading of the argument is it seems to presume that assent to a judgment cannot be *both* naturally caused *and* based on rational grounds, but without further argument it is not clear why the fatalist should accept this.⁶⁷

67. For interpretations along these lines see Rosefeldt (2000: 169–84), Wolt (2018: 188–

However, as I have presented Kant's conception of a free intellectual faculty it is clear that Kant must find the fatalist's attempt at argument incoherent. For it isn't (or isn't just) that if the fatalist is correct one's judgments would be, at best, epistemically lucky (and thus unreliable). Rather, it is that according to the fatalist the judging subject would be, as Kant puts it, a mere "plaything" (*ein Spiel sein*) of her sensible impulses. In acting according to subjective grounds—that is, empirical natural (psychological) causes—she would not be acting in accordance with rational intellectual laws. Thus the fatalist would not be acting spontaneously but merely receptively and would thereby lack any imputability with regard to her acts. Hence, she could not be seen as engaging in anything like an argument concerning what is true. Kant holds that any *argument* the fatalist makes presumes a use of a free intellectual faculty—that is, a faculty whose causality is not simply that of temporally determined nature, and works instead entirely according to its own laws. The truth of fatalism would thus undermine any possibility of arguing for its truth, and in that sense any self-conscious attempt to argue for fatalism is self-defeating.⁶⁸

There is a further issue that Kant's conception of the constitutive causal laws of a faculty raises. What evidence do we have for thinking that we possess rational faculties that could be the ultimate causal sources of our mental states? Do we have any good reason to think that we, rather than the mechanism of nature, are the ultimate causal source of our thoughts? The constitutive laws argument, like the others above, tells us nothing about our status as rational beings; it merely indicates that rational beings would have to satisfy the conditions it outlines. In the next and final subsection I look at one argument Kant offers in favor of a substantive positive answer to the question of our status as rational beings. The view that emerges from his work, I argue, is one according to which we have at least *some* evidence for construing ourselves as genuine thinkers, but this evidence cannot ever rise to the level of proper objective cognition, and so cannot be construed as servicing a genuine form of positive scientific knowledge concerning our rational status as thinkers.

89ff); see also Paton (1947: 218), Jordan (1969: 53ff), Allison (1996b: 63–64 and 99–100), Wood (1999: 177–78; 2008: 131–32). Allison contrasts "arriving at one's belief that p by the correct causal route," with "taking or recognizing these prior beliefs as warranting the belief that p" (1996b: 63). This is, in my view, somewhat infelicitous. If the exercise of one's intellect is the actualization of a causal power, then it would be better to say that one cannot arrive at a belief via the correct causal route (e.g., in an act of reasoning) unless one has, in an appropriate fashion, taken one's state as based on another.

68. So as I understand the argument, Kant isn't (or isn't just) pointing out a form of practical incoherence in the fatalist's position but rather indicates that it is straightforwardly theoretically incoherent to *argue* for fatalism—i.e., the metaphysical conditions for so acting (i.e., making an argument in which judgments are rationally held as true) cannot be satisfied; cf. Wood (2008: 132).

4.4. *The Argument from Self-Consciousness*

In the Transcendental Deduction Kant famously says that,

The **I think** must **be able** to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible or else at least would be nothing for me. (Transcendental Deduction §16, B131–32)

This text is typically interpreted as saying that any representation whose content could be something *for* the subject (i.e., present some possible, truth-evaluable state of affairs to the subject) must be such as to be self-ascribable. This is to read Kant as arguing only that rational subjects must be such as to possess the capacity for self-consciousness of the content of states that they undergo or to which they are otherwise subject. But Kant could plausibly mean something more in this passage than merely that in any act of thinking one must be able to become aware of oneself as a subject of that act.

In lecture notes Kant is reported as having argued for consciousness of the causal activity of the thinker in acts of thinking. For example,

When I say: I think, I act, etc., then either the word I is applied falsely, or I am free. Were I not free, then I could not say: I do it, but rather I would have to say: I feel in me a desire to do, which someone has aroused in me. But when I say: I do it, that means spontaneity in the transcendental sense *<in sensu transcendentali>*. But now I am conscious to myself that I can say: I do; therefore I am conscious of no determination in me, and thus I act absolutely freely. (*Metaphysik L₁* 28:268 (1777–80))

Kant distinguishes here between two senses in which one might be self-conscious. One sense might be described as a kind of reporting—viz. the sense in which one is aware of oneself as a subject of representation, and thereby self-ascribes that representation. This is the sense that Kant is typically construed as having in mind in the Deduction passage quoted above. But a further notion used in the text is that of the subject as *cause* of the thought, or of thinker as the agent of thought. This “agential” sense is also at work in Kant’s reported rejection of Spinozism.

when I think, I am conscious that my I, and not some other thing, thinks in me. Thus I infer that this thinking in me does not inhere in another thing external to me but in myself, and consequently also that I am a substance, i.e. that I exist for myself, without being the predicate of another thing. (*Religion Pölitz* 28:1042 (1783/84))

The point Kant is described as making in this passage concerns the manner in which the sort of self-consciousness characteristic of thinking—what Kant calls “pure” as opposed to “empirical” apperception—is a consciousness of oneself as the causal source of one’s thought, rather than of something else bringing that representation about *in* one. Kant also links here, as he does in the Analogies, causal activity and substantiality (A204/B250).⁶⁹ This link is specifically via the notion of being the ultimate causal progenitor of the thought. As Kant says above, “when I think, I am conscious that my I, and not some other thing” is what is doing the thinking.

In my view, Kant links the subjective and the agential sense of “I think” in a conditional manner. One cannot be non-empirically aware of oneself as the *subject* of a thought unless one is aware of oneself as the *agent* or *cause* of the thought.⁷⁰ One can of course be empirically conscious of mental states for which the agential

69. For discussion of the issue of one’s consciousness of one’s status as a substance see Wuerth (2014); McLear (a).

70. The subject/agent distinction has gained some prominence in discussions of the phenomenon known as “thought insertion”—the having of thoughts from which one is alienated or otherwise disowns—in contemporary psychology and philosophy of mind. The distinction is championed in Stephens and Graham (2000) and subsequently taken up by a variety of philosophers and psychologists. See, e.g., Radden (1998), Gallagher (2000), Coliva (2002), Bayne (2004), Kriegel (2004: 202 note 10), Duncan (2019). If I am right about Kant’s emphasis on the import of agent awareness for understanding “pure” self-consciousness then any account characterizing the nature of self-consciousness in some mediated manner is going to be mistaken. For example, Watkins (2005: 278–82) argues that pure self-consciousness provides only an *indirect* awareness of the self, via—presumably “direct”—awareness of activity and/or its results. This strikes me as problematic. For one, it does not account for texts such as (B429), where Kant says that “in the consciousness of myself in mere thinking I am the being itself”—Watkins’s interpretation seems to require construing “in” here as “via” and thus as grounding a kind of inference. But Kant does not say here that there is consciousness of activity from which one infers that it is oneself of which one is aware as so acting. Such an indirect route to self-consciousness is also incompatible with the *Pölitiz* lecture’s claim that when “when I think, I am conscious that my I, and not some other thing, thinks in me.” Whether or not Kant is subsequently correct in that lecture that one can infer from one’s consciousness of agency to the conclusion that one is a substance, it seems clear from the quoted passage that he regards the act of thinking as requiring that one be able to be conscious of oneself as the agent or progenitor of the thought. Further, an interpretation like Watkins’s leaves open the following question: how is it that one is conscious of the activity by which one becomes “indirectly” aware of oneself as *one’s own activity*? In fact, given Kant’s skepticism regarding inferences from effect to cause (e.g., A368) it will always be uncertain whether any given instance of activity awareness will license an inference to oneself as the agent of such activity. Indeed, it is hard to see how one could *ever*, from such materials, securely arrive at the conclusion that it is oneself who is the agent. I take Kant’s view to be exactly the reverse of what Watkins suggests. One is conscious (in pure apperception) of the intellectual activity in oneself (i.e., in the terminology introduced above one is conscious of oneself as the subject of the activity) only in virtue of one’s consciousness of oneself as the *agent* of the activity—i.e., as producing through one’s act a determinate effect.

condition is not satisfied; for example, in the empirical consciousness of inner and outer sense.⁷¹ But what is important for cognition, as Kant notes in the “*cogito*” claim at B131–32, is the *thinkability* of one’s representations. Intuition is “given prior to all thinking” (B132), but if that intuition is to be something *for* the thinker (i.e., to play a role in cognition via its incorporation in experience) rather than to merely be *in* the thinker (and thus play at best a merely natural causal-mechanical role in determining the thinker’s actions), then it must be possible for the intuition to be taken up by the unity of apperception, and thus be a part of the intellectual activity of the cognizing subject.

I thus take Kant to have material for the following argument.

1. One cannot use <I> correctly in acts of conscious thought unless one has control of that thought.
2. One has control of a thought just in case the thought’s occurrence is due to the exercise of one’s intellectual capacities according to their constitutive laws.
3. The occurrence of a thought due to the exercise of an intellectual capacity is incompatible with its occurrence being due to the mechanism of nature.
4. ∴ One cannot correctly use <I> in thinking unless one’s thoughts occur due to intellectual causes distinct from the mechanism of nature—that is, unless one’s acts of thought are transcendently free.

This argument hinges on the conceptual link between correct use of <I> and transcendental freedom. However, the claim that first person thought must be transcendently free generates a clear interpretive problem. One might reasonably claim on the basis of this argument that at least from one’s own case one can *know*, or at least *cognize*, one’s transcendental freedom. This is because if the above argument is sound, then from one further assumption—viz. that anyone who can consider the truth of this argument is correctly able to use the first-person concept—knowledge of transcendental freedom is secured. Kant of course denies in the first *Critique* that we have any such knowledge of transcendental freedom (A558/B586). So mustn’t Kant reject the argument? I think Kant’s considered position is that the argument is sound. What he rejects is the further assumption of theoretical knowledge of one’s correct use of <I>.⁷² We cannot have theoretical

71. One might also be conscious of various sensory imagery that typically accompanies episodes of thinking, such as verbal (i.e., “inner speech”) or visual imagery. But such imagery should not be mistaken for thought, and one might have such imagery without having the corresponding awareness of oneself as agent. Indeed, cases of thought insertion might be understood as cases in which there is verbal imagery without accompanying agential awareness of one’s producing the thought that typically accompanies such imagery. This assumes, however, that one does not equate or closely identify inner speech with conscious thought as with Ryle (2009), Carruthers (2011), Byrne (2018).

72. Note that this conception of competence with the first-person concept goes beyond the

cognition of the fact that we are in control of our thoughts, so we cannot have theoretical cognition of whether we are in fact correct users of the first-person concept. Kant's position on this point is not always consistently expressed.

For example, in the Clarification of the Cosmological Idea of Freedom Kant says,

the human being, who is otherwise acquainted with the whole of nature solely through sense, cognizes himself [*erkennt sich selbst*] also through pure apperception, and indeed in actions and inner determinations which cannot be counted at all among impressions of sense; he obviously is in one part phenomenon, but in another part, namely in regard to certain faculties, he is a merely intelligible object, because the actions of this object cannot at all be ascribed to the receptivity of sensibility. We call these faculties understanding and reason. (A546–47/B574–75)

The claim that one has cognition of one's intellectual acts in pure apperception is rather problematic and we can get clear on why this is the case once we understand Kant's conception of cognition. Kant uses the term "cognition" (*Erkenntnis*) in different ways.⁷³ For example, there is his very general definition of cognition as a conscious representation of an object (A320/B376). In this sense both intuitions and concepts may be cognitions. But even in this broad sense the claim that we cognize our free intellectual activity is problematic for there is no object given in pure apperception.

Kant also has a narrower sense of "cognition"—cognition in the "proper sense"—which, in finite beings such as ourselves, only concerns those states that are the outcome of a synthesis of sensory states with concepts.

to bring this synthesis to concepts is a function that pertains to the understanding, and by means of which it first provides cognition in the proper sense. (A78/B103; see also *Real Progress* 20:273 (1793))

I think we can see this narrower sense concerning the relation between concepts/judgment and intuition as motivated by a set of commitments that Kant endorses with respect to cognition—viz. (i) consciousness, (ii) agreement/conformability, (iii) real possibility.⁷⁴ Let me first briefly remark on each of these conditions before

kind of minimal competence outlined in the "fundamental reference rule" for <I>, such that "in using 'I' in the argument-place of her judgment, the judging subject is representing the fact that she, the judging subject, is the entity of which the predicate of her judgment is asserted to be true" (Longuenesse 2017: 23). This version of the reference rule is compatible with the judging subject not having "control" over her thoughts in the manner I specified above. For related discussion of self-reference, self-knowledge and causal ultimacy see McLear (2019).

73. The notion of "cognition" (*Erkenntnis*) has itself come under significant scrutiny of late. See Schafer (in press-a; in press-b), Willaschek and Watkins (2017), Watkins and Willaschek (2017) for representative discussion.

74. I'm indebted to Karl Schafer and Nick Stang for discussion of these issues.

going on to show how pure self-consciousness fails to rise to the level of cognition proper.

The consciousness condition is clearly stated in Kant's *Stufenleiter*. Cognition must be a conscious relation to an object. The conformability condition is also relatively straightforward. Kant conceives of cognition in terms of the "agreement" or "conformity" (*Übereinstimmung*) of a cognition with its object. For example, in discussing truth in the first *Critique* he says that the "nominal definition of truth, namely that it is the agreement of cognition with its object, is here granted and presupposed" (A58/B82). Since Kant allows for false cognition, I take his view to be that representations that are cognitions are such as *to be able* to agree or conform with their objects, and Kant construes such agreement as truth. So cognition need not be actually true to stand as cognition, though it must be such as to be possibly so. Moreover, Kant regards judgment (i.e., a specific sort of non-associative unity of concepts) as the bearer of truth, while he denies that intuition is the sort of thing that can be true or false.⁷⁵

Finally, Kant construes cognition as always needing to be of a really—that is, metaphysically—possible subject matter. As Kant states in a famous footnote in the B-preface of the first *Critique*:

To *cognize* an object, it is required that I be able to prove its possibility (whether by the testimony of experience from its actuality or *a priori* through reason). But 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. But in order to ascribe objective validity to such a concept (real possibility, for the first sort of possibility was merely logical) something more is required. (Bxxvin.)

In finite discursive beings the structure of thought is governed ultimately by the rule of non-contradiction, and in this way concerns what is logically possible. But an analysis of the contents and laws of thought tells us nothing as to whether there really could be an object such as the thought specifies. Thus there can be merely discursively or logically possible represented objects (e.g., a two-sided plane figure) that are not really possible. There may also be really possible objects that are not representable discursively (e.g., the world-whole as a "synthetic universal"). Thus the real possibility of an object for theoretical cognition can only be determined from an actual ground provided via experience or via a priori reflection.⁷⁶ And as Kant goes on to note, the "more" that is required for a thought

75. For discussion and defense of the claim concerning judgment see Heis (2013: 277–78); for intuition see McLear (2016).

76. Kant also distinguishes cognition from mere "acquaintance" (*Kenntnis*) and connects the

to achieve cognition need not always come from theoretical sources, as there may be practical sources of cognition as well.⁷⁷

We can thus take “cognition in the proper sense,” at least in its theoretical guise, as necessarily (i) conscious, (ii) truth-apt (i.e., discursively conformable), and (iii) of a really possible subject-matter. Kant’s conception of cognition in such terms explains why he takes satisfaction of these conditions in finite rational beings (i.e., cognition in its “proper” sense) to require *both* intuition and concept, since discursive beings are such that they cannot satisfy all three conditions without the integrated contribution of both sorts of representation.

Kant’s critical notion of (theoretical) cognition in finite discursive beings thus requires a representational complex in which judgment and intuition are suitably related via an act of synthesis. The critical Kant is clear that one cannot have an intuition of the mental act by which thinking occurs.⁷⁸ Moreover, one cannot have cognition in the proper sense via pure apperception alone, since there is no (and could be no) corresponding intuition with which to synthesize the appropriate judgment concerning one’s intellectual activity. So I take it that Kant’s considered view must be that self-consciousness cannot ever provide *cognition* that one is the agent of one’s mental states. This also helps Kant avoid the objection that pure self-consciousness provides a sufficient basis for positively *knowing* that one is transcendently free.⁷⁹ Such positive knowledge would require that one has cognition of one’s transcendently free acts. But of such acts one can *never* have cognition in the proper sense; they are at best acts of which one is merely *conscious*.

We can see from this account, however, why Kant might be inclined in various places to nevertheless construe pure self-consciousness of one’s mental activity as a sort of cognition, for even in the second edition of the Paralogisms Kant

difference between cognition and acquaintance to the consciousness of a ground of difference. See FS 2:59–60; B414–15; JL 9:64–65; *Dohna-Wundlacken Logik* 24:730–31 (c. 1792); *Wiener Logik* 24:846 (1780–82); *Blomberg Logik* 24:132–33, 134–35, 136 (c. 1771). For further discussion see McLear (in press; b).

77. For some discussion of the issues involved in practical cognition see Chignell (2007a), Kain (2010), Schafer (in press-b).

78. Inner intuition can only grant a *passive* awareness of something that happens or occurs in one’s mind—viz. some sensory event or episode. It cannot represent the *making happen* that one engages in as a thinker. However, Kant’s view on this matter in the run-up to the critical period is an altogether different story. In the 1770 *Dissertation* Kant speaks of attending to acts of the mind in experience (ID §8, 2:395; see also ID Corollary, 2:406 and ID §23, 2:411). And in lecture materials throughout the 70s Kant claims that we have intuitions of the self as it really is, in contrast to our intuition of external phenomena. See, e.g., *Anthropologie Collins* 25:15 (1772–73); *Metaphysik L₁* 28:206–207, 224–25 (1777–80). For relevant discussion see Mohr (1991: 118–38; 1995), Klemme (1999).

79. Detailing the complex relationship between cognition and knowledge requires more space than I can devote to it here. For some relevant discussion see Chignell (2014), Watkins and Willaschek (2017).

maintains that “in the consciousness of myself in mere thinking I am the being itself” (B429) that is thinking, which is to say that in pure apperception I am non-empirically aware of my existence, and of the acts of thinking that I, as existing, bring about. This self-consciousness is both a form of consciousness and of an actual being (the thinker), and thus satisfies two of the three necessary conditions on cognition outlined above—viz. the consciousness constraint and the real possibility constraint.⁸⁰ What fails to be clearly satisfied is the conformity/truth-aptness constraint, for one of the overall morals of the Paralogisms is that pure apperception does not provide sufficiently determinate material for reasoning about oneself, and thus nothing which one could articulate as a categorial claim concerning oneself as object.⁸¹ Indeed, Kant goes so far as to say that,

Thinking, taken in itself [*für sich*], is merely the logical function and hence the sheer spontaneity of combining the manifold of a merely possible intuition; and in no way does it present the subject of consciousness as appearance. . . . In this way I represent myself to myself neither as I am nor as I appear to myself, but rather I think myself only as I do every object in general from whose kind of intuition I abstract. If here I represent myself as **subject** of a thought or even as **ground** of thinking, then these ways of representing do not signify the categories of substance or cause, for these categories are those functions of thinking (of judging) applied to our sensible intuition, which would obviously be demanded if I wanted to **cognize** myself. (B428–29)

Kant here denies that pure apperception presents any sort of sensory or intellectual appearance. Moreover, the “being itself” of which one is conscious in being engaged in intellectual activity is not itself represented in terms of an object. This is what I take Kant to mean in saying that the concepts <subject> and <ground> do not “signify the categories” <substance> and <cause>. I take Kant to be saying that one is aware in pure apperception of being the subject and ground of a thought, but that this awareness is not of the kind necessary to yield

80. There is some room for debate concerning satisfaction of the real possibility constraint since the form of consciousness we have of ourselves is not intuitive or experiential—Kant denies that we have any intuition, either sensible or intellectual, of ourselves as active beings. However, depending on what Kant means in his note at Bxxvi by “a priori through reason,” the self-consciousness of the rational being in acts of thinking may qualify as possessed “through reason” or the use of the intellect more broadly. This seems especially true when considering Kant’s use of “a priori” in its more archaic “from-grounds” sense (Adams 1994: 109; Smit 2009; Hogan 2009: 53–34; Stang 2019: 98–101), denoting that through which something is explained. Since one’s acts of thinking are explained by the thinker and not anything else, the actuality of the thinker of which one is or must be able to be conscious is proved a priori in any intellectual act, and in this sense “through reason.”

81. For a similar position see Wuerth (2014: 144–45).

cognition, for the application of <subject> or <ground> here, and the related categories, is not warranted by any intuition. Thus there is a sense in which it is incorrect to construe the content of one's thought here as constituted in terms of the application of <substance> or <cause>.

One might thus wonder in what sense the consciousness of oneself as thinker has any content at all, if it has no categorial content concerning a substance or its causal activity. Here I take Kant to construe the content of such self-consciousness as not including <substance> or <cause> even if it does contain the closely related <subject> and <ground> because of two reasons. First, the content of such self-consciousness is unschematized (it is not of a possible appearance). Second, lacking intuition of oneself in thought, there is no "object"—in Kant's demanding and technical sense—presented to consciousness. Kant tends to use "content" (*Inhalt*) to mean "relation to an object," and there is no such relation here. But this fact is compatible with our broader use of the term in contemporary philosophy, according to which "content" means something more like "information conveyed."⁸² Given that wider contemporary sense of "content," Kant can consistently deny that pure self-consciousness provides any relation to an object, and thus material for judgment, without being forced to say that pure self-consciousness fails to convey anything at all to a subject.

Kant also acknowledges that how things seem in the consciousness of intellectual activity might not be the case. For example, while discussing the "Canon problem" in Section 3.1 we saw that Kant at least entertains the possibility that the activity of reason of which one is conscious might ultimately have its causal source in nature—in "higher and more remote operating causes" (A808/B831)—rather than the free faculty of reason itself.⁸³ So Kant allows that while it certainly seems to us as if we are engaging in rational acts such as inference, it is at least epistemically possible that we are not doing so. Thus we cannot know for certain that our reason is not ultimately determined by alien influences, and so cannot know for certain that we are genuinely thinking and reasoning, rather than engaging in their mere *simulacra*.⁸⁴

However, despite our ultimate ignorance concerning our status as rational beings, Kant can also argue that we have sufficient grounds for the legitimate *belief*

82. See Tolley (2011: §5) for the point about Kant's use of "content" (*Inhalt*). For a useful overview discussing the contemporary notion of "content," though with a focus on perceptual experience, see Siegel (2010).

83. See also Ameriks's (2000a: 69–70) discussion of the possibility that one's acts of first-person thinking are the result of realization by multiple substances, though the issue could just as well be pressed by appeal to a single substance (such as God) of which we are all accidents. This is also suggested by the fact that Kant claims in the First Paralogism that "it is not determined whether I could exist and be thought of only as subject and not as predicate of another thing" (B419).

84. See also Allison (1996b: 127).

(*Vernunftglaube*) that we are such beings. Indeed, the claim that we are rational beings capable of being the agents behind our acts of attending, conceptualizing, thinking/judging, etc., is a necessary precondition or postulate of reasoning at all, on analogy to Kant's conception of justified moral belief in the existence of God or of our own immortal souls.⁸⁵ This sort of "doctrinal belief", as Kant calls it (A825–26/B853–54), is a theoretical analogue of practical belief, and concerns that for which we can never have sufficient objective grounds for holding true, but which is something to which we are nevertheless rationally driven.⁸⁶ Moreover, this drive is distinctively theoretical rather than practical, concerning the interest of reason in what *is* rather than what *ought* to be (CPrR 5:120; A804/B832).⁸⁷ Following Stang (2016a: 263) we can call the kind of belief at issue here "necessary theoretical belief," which is belief in a postulate that is necessary to satisfy theoretical reason's need for explanatory unity, but whose truth cannot be cognized or known by us. As Stang (2016a: 287) points out, such necessary theoretical belief, if it is to be relevantly analogous to moral belief, must be of something that is knowably logically possible and must not knowably be really impossible.⁸⁸ Our status as rational causes of our intellectual acts satisfies this, at least insofar as the argument that we can be progenitors of new causal chains, as set out in the Third Antinomy, is cogent.⁸⁹

85. There is dispute in the secondary literature as to whether Kant has a conception of rational belief outside that of moral belief. For defense of the strong coincidence of rational with moral belief see Wood (1970), Stevenson (2003), Pasternack (2011; 2014). For defense of there being more kinds of rational belief than just moral belief see Chignell (2007a), Santi.

86. On this point see especially Chignell (2007a: 345–54).

87. For discussion of practical vs. theoretical grounds see Chignell (2007b), Kain (2010), Schafer (in press-b).

88. This sets my proposal apart from a view such as Korsgaard (1989). Korsgaard construes the structure of first-person deliberation as indifferent to the issue of freedom or determinism. As she puts it, "Kant's answer to the question whether it matters if we are in fact (theoretically) free *is* that it does not matter" (1989: 40). Regardless of the theoretical truth or falsity of determinism, one must deliberate *as if* one were free. This broadly compatibilist reading of Kant is also displayed in her *Sources of Normativity*, where she states that "Determinism is no threat to freedom" (1996: 95). Korsgaard construes the issue of freedom relevant to normativity to be entirely practical rather than theoretical (1996: 96). While Korsgaard is certainly right that even if one possessed theoretical knowledge of the truth of determinism one would still need to go through the activity of deliberation (1996: 94–97), nevertheless if determinism (or indeterminism) were the whole truth, it would leave us as the mere turnspits Kant decries in the second *Critique*, mechanically acting out a process of deliberation while lacking the causal ultimacy to make that deliberation relevant to imputation. Thus, we could not have belief in our freedom in either of its theoretical or practical senses if it were knowably really impossible for us to realize it; see also Allison (1996b: 127–28).

89. This does mean that the discussion and ultimate resolution of the Third Antinomy is absolutely central to Kant's entire critical project, more so than might seem indicated by its position in the architectonic of Kant's first *Critique*. However, I think we should not confuse the position an argument holds in the architectonic with the relative position (or centrality) of an

Consciousness of our representational activity via pure apperception thus provides at least some reason for thinking, as the constitutive argument demands, that we have rational faculties and thus could be the ultimate causal sources of our intellectual acts. But we can never, at least on theoretical merits, *know* or even *cognize* that we have such faculties, and thus know or cognize that we engage in such transcendently free acts. Nevertheless—in analogy with the moral case—it seems that we have a legitimate basis for belief (*Glaube*) in our intellectual freedom.

5. Two Objections

The position I've outlined is no doubt controversial. Here I discuss two general objections that one might raise. These are, first, that the intellect, at least in its purely theoretical use, might be merely “relatively” spontaneous rather than transcendently spontaneous; second, that Kant revises his commitment to the freedom of intellect after the mid-1780s because of a supposed change in his strategy in arguing for freedom. I take these in turn.

5.1. Relative Spontaneity

Perhaps the most well-known objection to the position that Kant conceives of the intellect as transcendently free is from Wilfred Sellars, who argues that,

we can conceive Kant to argue that although we are conscious of ourselves as *spontaneous* in the synthesizing of empirical objects, this spontaneity is still only a *relative* spontaneity, a spontaneity “set in motion” by “foreign causes”. (1971: 23)⁹⁰

And slightly later,

Kant is leaving open the possibility that the being which thinks might be something “which is not capable of imputation”. It might, in other words, be an *automaton spirituale* or *cogitans*, a thinking mechanism. (1971: 25)

Sellars is arguing here for two points. First, that Kant distinguishes between an “absolute” spontaneity of transcendental freedom and a merely “relative” spontaneity, that is appropriate only to a “mechanism.” Second, that Kant allows for the possibility that the theoretical use of the intellect in the cognition of objects is merely relatively spontaneous.

argument to Kant's critical philosophy. For defense of the centrality of the Third Antinomy for Kant's overall critical position see Willaschek (2006).

90. For critical discussion see Pippin (1987: §4), Kohl (2015b).

We've seen, in Section 3.2 above, that Kant rejects the kind of temporally determined relative spontaneity proffered by a broadly Leibnizian model. Let's assume that Sellars doesn't have this kind of position in mind. But then what exactly is the sense in which the theoretical use of reason evinces a spontaneity that is merely "relative," such that it doesn't allow for imputation, in contrast with the practical use of reason, which does?

The main problem, as Sellars conceives of it, is that the theoretical use of one's intellectual faculties might work according to principles, while nevertheless lacking any purpose or norm-providing end towards which one's acts strive.⁹¹ The way that Sellars puts this difference between the theoretical and practical uses of reason is in terms of the presence in the latter use of what he calls a "practical premise":

for practical reason to be autonomous, there must be a practical premise which is as intrinsic to reason as are its conceptual or "transformational" procedures. For, surely, if all its premises come from without, then it is indeed "set in motion" from without—its "causality" is "caused"; its "spontaneity" relative. (1971: 26)

The implication here is that the theoretical use of reason is not governed by a "premise" intrinsic to it, rendering it merely relatively spontaneous. The "premise" is supposed to be something more than a simple "transformational" rule or principle, rather it must be something that articulates a purposeful use of or end for the activity of the faculty. Sellars ultimately attributes such purposeful activity only to the practical intellect. He says,

it is clear that although the structure of the first *Critique* highlights what I have called the relative spontaneity of the conceptualizing mind, it clearly presupposes a larger context in which the mind is thinking to some purpose. Thus reference to reason in its practical aspect is implicit throughout the *Critique*, but only in the *Dialectic*, after the constructive argument is over, does it become explicit. (1971: 25–26)

Sellars argues that while, for Kant, self-love is a purpose "implanted in practical reason from without" the moral law counts as an internal purpose or "intrinsic practical premise" (1971: 27). Interpreted this way, Sellars argues that the practical use of one's faculties involves an intrinsic purpose for their use that is not present in the theoretical use of one's faculties, and for that reason the theoretical use of one's intellectual faculties is merely transformational (i.e., "relative") and not genuinely (or "absolutely") spontaneous.

⁹¹. Thanks to Yoon Choi for helpful discussion of Sellars's argument.

But positing such a disparity isn't cogent. Most importantly, Sellars draws a thoroughly false contrast between the merely "transformational" procedural acts of the theoretical use of reason and the "intrinsically purposeful" acts of practical reason.

As I discussed in Section 4.3, Kant distinguishes the two uses of reason according to the distinct intrinsic aims that their use accomplishes, each of which cannot be reduced to the other.

The interest of its speculative use consists in the cognition of the object up to the highest a priori principles; that of its practical use consists in the determination of the will with respect to the final and complete end. (CPrR 5:120; see also A666/B694; A686/B714; A797–804/B825–32, *passim*)

The interest of theoretical reason consists in cognition of what *is*, in accordance with "the highest" of a priori principles provided through the unifying and systematizing activity of theoretical reason. The interest of practical reason lies in the determination of the will, again from a priori principles, concerning what *ought* to be, in the sense of realizing the "final and complete end" of the highest good. But because each use of the intellect has an intrinsic aim or end, exercises of that faculty are themselves governed by that end. For example, one cannot engage in acts according to the interests of theoretical reason if those acts are not in accord with laws that are or are suitably related to logical laws.⁹² Sellars is thus wrong to conflate the interest of reason in its "practical aspect" with the interests of reason *simpliciter*, and he is wrong to think there is a special disparity to be found between the intrinsic purposes of the practical as opposed to theoretical use of reason (or the intellect more broadly).

If Sellars's argument as stated cannot drive a wedge between the theoretical and practical uses of reason (or the intellect more broadly) based on appeal to purposes, might there not be other ways? Might the theoretical use need "material from without" in a way that the practical does not? Or might the intrinsic principles governing the practical be more clearly specifiable than those governing the theoretical? Or, finally, might the practical use of reason be "spontaneous" in the sense of providing material for reasoning that is novel by going beyond any possible sensory input? I believe Kant can give straightforwardly negative answers to all three of these attempts to differentiate the spontaneity of the practical from the theoretical use of reason.

First, practical reason, just as with its theoretical use, requires material upon which to work, so there is no disparity between their characteristic forms of activity to be found at the point of what initially causes reason to spring into

92. See also my discussion of Kitcher in Section 4.1.

action—in both cases some input from sensibility is necessary.⁹³ Hence, if we take Sellars’s initial gloss of a relative spontaneity at face value, as that which is “set in motion by foreign causes” it is unclear that there is any difference between the practical and the theoretical elements of the intellect. *All* of one’s intellectual faculties ultimately depend on being first set in motion by the deliverances of sensibility.⁹⁴

Second, as we saw in the discussion of the constitutive laws argument, the intellect must have principles intrinsic to its activity in order for it to be individuated as a distinctive faculty or power for the generation of representations. Such principles presumably include the principle of contradiction, which Kant says is the “supreme principle of all analytic judgments” (A150/B189). Since general logic concerns analytic relations (e.g., A76/B102) and the “absolutely necessary rules of thinking, without which no use of the understanding takes place” (A52/B76), the principle of contradiction is clearly an intrinsic principle governing the operation of the intellect as such. It seems plausible that similar considerations can be raised for the operation of reason and the form of the principle of sufficient reason Kant articulates as the principle driving reason’s search for the conditions of any given conditioned thing (A307/B364).⁹⁵ So the theoretical use of the intellect has clearly articulable intrinsic principles and premises just as the practical use does.

Third, it is part of Kant’s conception of *theoretical* reason (here in its narrow sense as faculty of inference and progenitor of ideas) that its activity is not exhaustively understood as a response to sensible material—“it does not follow the order of things as they are presented in intuition, but with complete spontaneity it makes its own order according to ideas” (A548/B576; see also GIII 4:452).

93. Kant does note important differences between sub-faculties of the intellect—e.g., understanding and reason—based in part on the manner in which those faculties generate concepts (i.e., reflection vs. inference). And there is a clear difference, as discussed in Section 4.1, concerning the way in which the will and power of choice are operative in the theoretical vs. practical forms of agency. But Kant can consistently hold that there are important differences between the various sub-faculties while holding that the intellect in general (i.e., in both its practical and theoretical uses) cannot engage in its various activities without some initial material from the senses.

94. Famously, Kant denies that there are any innate representations in the mind. So while the *receptivity* of the subject is innate, the actual sensible nature of that receptivity (viz. sensibility as specifically spatial and temporal) is acquired in what Kant calls an “original acquisition” (OD 8:221–22). Similarly, one’s capacity for spontaneity is innate, but Kant speaks of the “givenness” (A728–29/B756–57) of the categories, as “reflected concepts” for which the subject’s experiences stand as the “occasional causes of their generation” (A86/B118). For discussion see Allison (1973), Vanzo (2018). See also Refl. 4172 (1769–1770), 17:443; *Metaphysik L₁*, 28:190, 233–234 (1777–80); *Metaphysik Mrongovius*, 29:762–763 (1782/3); *Metaphysik Volckmann*, 28:373–374 (1784/85).

95. For related discussion and a similar approach to critiquing Sellars on this point see Kohl (2015b: §4).

Theoretical reason is thus not merely reactive to sensory stimuli in activities of reflection and synthesis, as is the understanding, but rather generates its own representations—ideas—by virtue of the principles governing its activity (specifically, the aforementioned principle of sufficient reason governing its search for systematicity and unity).

Sellars might concede that theoretical reason has its own constitutive principles, and that its actions outstrip what is given to it by sense so that it is not a merely reactive faculty, but nevertheless deny that possession of such constitutive principles of activity is sufficient for the kind of cognitive control and absolute spontaneity for which I have been arguing.⁹⁶ Sellars suggests that the relative spontaneity of the theoretical use of the intellect might be like that of a computer realizing “set dispositions” (1971: 25) in response to an external stimulus. I’ve advanced a conception of “cognitive control” in terms of acting according to the constitutive laws of one’s cognitive faculties without being determined to do so by an alien ground or cause. Isn’t this just action in accordance with “set dispositions,” as Sellars describes? If so then how is it an act of absolute spontaneity—that is, transcendental freedom?

It cannot be action from “set dispositions” *per se* that is the problem here, for God’s intellect works from set dispositions—laws from which he never deviates—and this fact does not show that God lacks absolute spontaneity. Moreover, one can object here to the very notion of describing intellectual capacities or abilities as “set dispositions” since <capacity> and <disposition> are not identical concepts. One can have a capacity without having any specific set disposition—that is, I have the ability or capacity to throw myself out a window, but I lack any disposition (settled or not) to do so.⁹⁷ In any case, the crucial idea is that intellectual acts occur according to specific, constitutive, laws. If Sellars is insisting here that an absolute spontaneity *doesn’t* operate according to set laws but rather must be able to deviate from them, then perhaps he is making one of two illicit assumptions.

On the one hand, Sellars might be operating with a conception of absolute spontaneity as (or as requiring) leeway, which I have already argued (in Section 3) that Kant rejected. Of course, there are ways in which one might exercise choice with respect to intellectual acts—one can choose what to attend to, what lines of inquiry to adopt or pursue, etc.—but one is not acting according to one’s rational powers if one refuses to accept q despite accepting p and *if p then q* , or if one tries to think a contradictory thought of the form $(p \ \& \ \neg p)$. In lacking the ability to do other than think a logically coherent thought or make an obvious logical inference based on premises to which one both attends and accepts one does not

96. Thanks to an anonymous referee for pressing this point.

97. See Vetter (2015: 19) for this example and discussion of the distinction between abilities and dispositions.

thereby forsake one's freedom/absolute spontaneity, rather one exemplifies that very freedom.

On the other hand, Sellars might instead assume that the causality with which one acts need not always be lawlike, but Kant clearly denies this.⁹⁸ And if Sellars's issue ultimately concerns neither of these points but rather that he holds that the theoretical use of the intellect is merely responsive in a manner that contrasts with the activity of its practical use, then we fall back on the various sorts of response I have given above.

Hence, *pace* Sellars, the theoretical use of the understanding and reason, just as with the practical use of reason, depends on the existence of principles constitutive of the causal activity of the faculty. Since these principles are rationally constitutive they will not count as "alien" influences stemming from sensibility or implanted by God.⁹⁹

There is one further issue about relative spontaneity, viz. that what "haunts" Kant is the possibility that we are merely spiritual automatons (Sellars 1971: 25). Recall from Section 4.4 Kant's argument in the *Metaphysik L₁* and *Metaphysik Pöhlitz* lectures that one cannot be non-empirically aware of oneself as the subject of a thought unless one is aware of oneself as the agent of the thought. Sellars's "haunting" question might then be put in terms of whether one could conceive of oneself as the merely *relative agent* of one's thought, so that all of one's self-conscious activity could ultimately be explained by mechanical causes beyond one's control.

Consider a mundane example. While talking with a friend, one has the spontaneous mental image of one's keys on the table at home (the image occurs to one "in a flash"), which then is the basis for one's judging that one left one's keys at home. This might seem an instance of "relative spontaneity" in one's thinking, governed by certain conditions from without, but proceeding according to internal rules. Why couldn't *all* thought be described in this manner?

98. See, e.g., Kant's definition of free causality in *Groundwork III* as that according to laws, for "freedom . . . must . . . be a causality in accordance with immutable laws but of a special kind; for otherwise a free will would be an absurdity" (G 4:446).

99. Note that it is one thing for God to *create* beings with intellectual or practical faculties. It is another thing to create beings with such faculties plus some internal "implanted" impulse that would cause them to act in specific ways. Kant is partly concerned with the latter point in his discussion of God's influence on created beings (e.g., CPPr 5:100). A similar point is discussed by Kant in his rejection of "preformationist" accounts of the categories (B167–68; see also Kohl 2015b: 319–20). In any case, the mere fact that one is *created* with a rational nature need not threaten one's transcendental freedom, for such freedom simply consists in having a particular faculty—viz. one whose causality operates according to specific, non-temporal, laws. As I argued in Section 3.2 above, Kant's worries about God's creative activity are primarily based on the worry that this would undermine the causality of our intellect. Kant also has worries about leeway, especially if time were not transcendently ideal, but these are downstream from the issue of absolute spontaneity and transcendental freedom.

While such episodes are colloquially described as episodes of “thinking,” Kant would consider that description contentious. At most, he would argue, the occurrence of the mental image itself is merely “relatively” spontaneous in the sense that it is generated by the imagination in response to stimuli. In contrast, the *thought* that one’s keys are at home, insofar as it is a rational judgment based on the content of the image, and not itself a mere happening, is something over which one has cognitive control. It is thus an act that one actively *does*, and in principle could refrain from doing (e.g., as when one knows one has been given a drug that will prompt the occurrence of random visual images). If the resulting state/activity *isn’t* a rational response to the image, but itself a kind of mere happening, Kant would deny that it is a judgment at all, being instead a kind of sensory association (as when one “hears” oneself “say” in an episode of internal speech, that one’s keys are at home).¹⁰⁰ In this latter case, what we would have is, at best, mere association, and not spontaneous thinking in the Kantian sense at issue. Moreover, the kind of mentation (to use a neutral word) involved in such mundane episodes as the example above describes is notably different in character from the experience of marshaling one’s thoughts in judgment or inference.

Thus Sellars’s view that thinking might be merely relatively spontaneous is simply contrary to Kant’s account of thought. If this is the possibility Sellars is suggesting bothers Kant, he would be wrong to allow that Kant is haunted by, or otherwise might endorse, such a view of the intellect. There *is* a related worry bothering Kant, but Sellars fails to adequately articulate it. What haunts Kant is the possibility that we may not be rational beings, our mental activity at best being a kind of simulacrum of a rational mind. He is thus not haunted by the possibility that we may be merely relatively spontaneous *thinkers* while being absolutely spontaneous *actors*, for there is no principled basis in Kant’s view for the relevant contrast. Instead the issue is whether we are genuine or merely ersatz rational agents. In Section 4.4 I argued that Kant has the resources to claim that we are *legitimately* driven to our belief that we are rational, and thus that we are genuine thinkers, though this belief can never be given a proof on purely theoretical grounds. But the principled basis by which Kant can argue for the legitimacy of the belief should be enough to exorcise the specter of Sellars’s “haunting” possibility.

Hence, in my view Sellars fails to provide for a substantive distinction between the form of spontaneity operative in practical as opposed to theoretical uses of the intellect as Kant conceives of it.¹⁰¹

100. A variety of empiricist conceptions of self-knowledge construe awareness of thinking in terms of awareness of episodes of internal speech. See, e.g., Ryle (2009), Carruthers (2011), Byrne (2018).

101. To be sure, there is an important difference between the operation of our faculty of choice and the operations of the intellect. But as I argued in Section 4.1, the freedom of the faculty of choice should not be straightforwardly identified with Kant’s concept of

5.2. Kant's Change of Strategy

As noted at the outset of this essay and often discussed in the secondary literature, Kant pursues a strategy for demonstrating our transcendental freedom in the early to mid-1780s that links together theoretical and practical spontaneity. This is reflected in Kant's 1783 *Review* concerning Johann Heinrich Schulz's work on ethics, discussed in Section 4.3. Kant's strategy for proving one's practical freedom through an appeal to the freedom of the intellect is taken to culminate in the *Groundwork III* argument for the autonomy of practical reason.¹⁰² However, Kant subsequently appears to discard this strategy, and perhaps even the entire aim of providing a demonstration of freedom, opting instead for the infamous "fact of reason" argument in the second *Critique*.¹⁰³

This raises the following worry: perhaps at one time (e.g., the 1770s) Kant accepted that the intellect is transcendently free, and pursued a strategy consonant with that position. But at least by the late 1780s his rejection of the strategy linking the theoretical and practical uses of reason might indicate that he no longer endorses the view that the intellect is free, at least in its theoretical use.¹⁰⁴ If that were correct, then either Kant's overall view is incoherent, or I am wrong about the conceptual connections between the nature of the intellectual faculty and freedom.

Against this worry, I argue that Kant maintains his view concerning the free intellect throughout the critical period, but subsequent to the *Groundwork* holds that cognition concerning the free activity of one's faculties can only be supported

transcendental freedom/absolute spontaneity. So while it is right that there are important differences between the activity of our theoretical as opposed to practical faculties, concerning the issue of transcendental freedom in particular, the nature of their operation is analogous.

102. Note that I am not presuming here that Kant's conception of (practical) autonomy is identical to the conception of control I've outlined here, though they must be closely related. However, I am assuming that an agent's will is autonomous in the sense discussed in the *Groundwork* just in case the agent has control of their acts of willing.

103. For discussion of these points see Paton (1947: 203–206 and chs. 20–24 *passim*), Henrich (1960; 1975), Prauss (1983: ch. C), Allison (1990: ch. 12), Wood (1999: 171–82), Ameriks (2000b: 191, 209–20; 2003a: ch. 6; 2012: 24–25), Kosch (2006: ch. 1), Guyer (2009), Tenenbaum (2012), Choi (2019). As Ware (2017: 117–18) notes, Kant's early interpreters did not see him as changing strategy in the second *Critique*.

104. Ameriks gives perhaps the most sophisticated version of this reading. While I find his discussion important and illuminating, I disagree with Ameriks over the extent to which Kant really wavers in his conception of the free use of the intellectual faculties. In my view this is an unwavering metaphysical commitment of Kant's, which goes back at least to the early 70s, and whose truth is necessary for the success of the entire critical system. To the extent that Kant takes an "elliptical path" to his later critical position, it is one concerning how to best situate the status of ourselves as rational beings within the critical epistemology. I construe the changes in presentation of the argument not so much as a refinement of "relatively crude beliefs" (Ameriks 2000b: 214) about freedom than as a clarification of the commitments of his theory of cognition and the limits of pure self-consciousness's contribution to cognition.

by practical as opposed to theoretical grounds. Hence, while Kant's conception of the epistemic grounds for accepting that we have such a free intellectual faculty evolves over the course of the critical period his basic conception of the nature of that faculty does not change. There are three considerations for favoring my preferred interpretation.

First, showing that the activity of the intellect even in its theoretical use, and not just its moral or practical use, presupposes transcendental freedom does not in and of itself provide us with overwhelming justification for holding that we can behave morally. The argument of *Groundwork III* in part consists of a "partners in crime" style argument that reason in its production of ideas necessary for theoretical cognition, exhibits a form of spontaneity that transcends all empirical conditioning. But a committed skeptical naturalist might deny this claim about theoretical reason just as thoroughly as any claim about the freedom of practical reason. It's not clear then, that Kant's argument in *Groundwork III* provides any real traction against a committed skeptical naturalist.¹⁰⁵ Indeed, acknowledgment of this fact all on its own, without requiring a change to Kant's underlying commitments, could have prompted him to change his approach in the second *Critique*, where his appeal to a "fact of reason" seems to be deliberately indifferent to skeptical worries.

Second, even if one accepts that the use of the intellect requires the transcendental freedom of that faculty, this does not show that the freedom of the faculty is such as to be able to determine the subject's power of choice (*Willkür*). For it is one thing to argue that a faculty can determine itself, it is another thing to argue that it can determine another faculty—that a rational law could be a law that governs choice.¹⁰⁶ There is no contradiction in conceiving of a being possessed of an intellectual faculty whose operations satisfy the conditions for transcendental freedom, but which lacks the power to determine its faculty of choice according to reason (e.g., perhaps because its choices can only be sensibly determined, or are wholly lawless). This point becomes especially salient once Kant clearly differentiates between will (practical reason; *Wille*) and choice (*Willkür*).¹⁰⁷

105. For this kind of worry see, e.g., Paton (1947: 244).

106. Kant makes this point quite clearly in his discussion of incentives in *Religion* 6:26n. Kitcher (2011: 246–47) takes Kant's discussion here as proof that theoretical and practical reason are importantly different with regard to freedom. This seems too quick. All the passage shows is that the theoretical use of reason, even if free in the sense for which I've argued, does not dictate how the power of choice is to operate, and in particular, whether it is ultimately inclined by purely intellectual considerations constituting a moral subject's "highest incentive", or heteronomous ones "coming from the objects of inclination". This is consistent with Kant's realization that the freedom of reason is one thing, and the freedom of the power of choice another, and thus the rejection of the strategy outlined in the *Review of Schulz* and *Groundwork III*.

107. Kant makes explicit the distinction between will and the power of choice in the *Metaphysics of Morals* (Introduction, 6:213), even going so far as to say that the will "cannot be

Third, since, as I argued in Section 4.4, pure self-consciousness of one's intellectual activity can never rise to cognition proper, any attempt to substantiate (theoretical) cognition or knowledge of one's freedom is going to fail. This means that there can be no demonstration of one's freedom based solely on premises appealing to theoretical grounds.¹⁰⁸

Thus, insofar as Kant's views in the 70s and the early-to-mid-80s concerning the extent of our theoretical cognition are in flux, he is thereby drawn to a view according to which the conscious awareness in pure apperception that one has of the self and its activity is enough for cognition (e.g., A546–47/B574–75, discussed above). But by the second *Critique* Kant more consistently recognizes that *consciousness* alone cannot provide for the cognition necessary to the pursuit of a demonstration of our freedom.¹⁰⁹ This more consistent position would not require Kant to change his mind concerning the connection between the intellect and transcendental freedom. But it would affect his pursuit of any argumentative strategy concerning the aim of demonstrating our transcendental freedom.

Given the above three points, it is perfectly consistent for Kant to continue to hold in the mid-to-late critical period that the intellect is transcendently free while coming to see that our status as rational agents cannot ultimately be proven on theoretical grounds, and that a shift to a proof on practical grounds is more consistent with the structure of the critical system as a whole. In this way the

called free or unfree. . . . Only *choice* can therefore be called *free*" (6:226). For discussion see Beck (1987), Hudson (1991), Kosch (2006: 60–61).

108. Alternatively, one might argue that Kant consistently distinguishes between a wider and a narrower sense of "cognition" such that he can consistently deny that theoretical cognition of the free intellect is possible in its narrow sense (what I have here called "cognition in its proper sense") while asserting the possibility of its cognition in the broad sense. See Keller (1998: 157), Kohl (in press) for defense of this kind of position. As far as I can tell such an argument is compatible with the position I defend here, though I take the central issue to concern not so much the verbal issue of whether there is a use of "cognition" in which one can cognize spontaneity, but rather whether our representation of the spontaneity of the intellect could ever satisfy all three of the conditions (viz. consciousness, truth-aptness, and real possibility) that I discussed above, the joint satisfaction of which Kant typically describes as cognition in its "proper" sense.

109. Though here one might appeal to Kant's discussion of the "the sole fact of pure reason" announcing "itself as originally lawgiving" (CPrR 5:31). Is there a fact of *theoretical* reason? Why does the issue only come up in the moral setting? Interestingly, in the CPrR Kant both says that it is reason alone from which rules that "contain necessity" can arise (5:20) and that we "can become aware of pure practical laws just as we are aware of pure theoretical principles, by attending to the necessity with which reason prescribes them to us and to the setting aside of all empirical conditions to which reason directs us. The concept of a pure will arises from the first, as consciousness of a pure understanding arises from the latter" (5:30). The "fact of reason" here is a consciousness of the necessity present in the operation of the intellect in *both* its theoretical and practical activity. Nevertheless, if cognition is to be had here it is still only *practical* cognition, for it is based only on the consciousness of a faculty insofar as it is called to use, and not in and of itself, as would be required in a theoretical proof of reason's freedom.

four arguments we've considered in Section 4 all outline the contours of a theory of rational activity without, on theoretical grounds, establishing that the human being satisfies such conditions. It is notable though that there do not seem to be considerations for showing that the human being *cannot* satisfy such conditions either.

6. Conclusion

In summary, Kant regards an act as the result of a substance exercising its causal powers according to specific laws. The intellect, as constituted by a specific set of causal powers, is thus a faculty for acting in specific ways, governed by intellectual laws, both practical and theoretical. The causality of intellectual acts differs from those of receptive acts of sensibility in the manner in which they incorporate their determining grounds, and thus are under the agent's control. The existence or nature of the determining ground of the act of a receptive capacity, which is not under the control of the agent, is independent of any exercise of that capacity. This is not the case for the spontaneous acts of the intellect. The causality of such intellectual acts includes not only instances of willing or intending, but also exercises of attention, the formation of concepts, the occurrence of judgments, the making of inferences, and the adoption or revision of one's doxastic attitudes. Such acts are transcendently free in the sense that the agent is the causally ultimate source of the actions, and her acts are not themselves determined by any factors beyond her control. Since the temporal causality of (phenomenal) nature is incompatible with the causal ultimacy and control of such acts, the status of intellectual acts as *rational*—that is, as occurring due to specific and distinctive laws—depends on the possibility of a non-natural (i.e., noumenal) causality.

Focusing primarily on the theoretical use of the intellect, I discussed four arguments elucidating the sense in which the intellect is transcendently free. I argued that, for Kant, the imputation of intellectual acts, and thus the possibility of making "ought" judgments concerning what an agent should, for example, believe or infer, depends on the transcendently free status of such acts. I then argued that transcendental freedom is assumed in any act of reasoning, since even cognitively basic acts—for example, the recognition of inferential connections between the content of one's representations—depends on a "taking" condition on reasoning that is incompatible with such taking being a product of a merely natural causality, and thus depends on reasoning being something under the subject's control. I subsequently argued that the plausibility of the claim that reasoning requires control over one's acts of reasoning is based on Kant's position regarding the constitutive (causal) laws of a faculty's activity. Finally, I argued that Kant adduces evidence for the existence in human beings of such causally operative rational faculties in the form of the "pure" self-consciousness one has of

one's intellectual activity in first-person judgment or reasoning. In such cases one is conscious that it is oneself that brings about one's thoughts and not some other force or cause acting through one. However, due to Kant's restricted conception of proper cognition, the "pure" first-personal consciousness of one's intellectual activity is insufficient for cognition and thus ultimately for positive theoretical knowledge that one is, in fact, transcendently free. But the claim that one can rationally act (at least sometimes), is something which one can legitimately believe, and thus does lend Kant some support for his other positions.

I thus hope to have elucidated the deep parallel between the rational activity of the intellect in its theoretical use and the rational will's practical activity in the formation of intentions to act. Proper appreciation of the way in which the concepts <substance>, <act>, <capacity>, <cause>, <freedom>, and <law> are intertwined show how deeply Kant is committed to a connection between freedom and rational activity. Though Kant's writings on normativity and freedom focus primarily on morality, as the human being's "highest end" (A801/B829), the activity of the intellect in its entirety, and not just with regard to moral action, must be the result of the agent's exercising cognitive control in expressing her power of absolute spontaneity, a power governed by constitutive laws distinct from those that govern phenomenal nature. There is, to be sure, much more that must be said regarding this view of rational agency. For example, is there the equivalent of one's moral character (*Gesinnung*) for the theoretical use of the intellect? Are there virtues and vices distinctive of the theoretical exercise of the intellectual faculty, or is there perhaps merely a broader notion of cognitive virtue of which the practical and theoretical are species?¹¹⁰

Whatever the ultimate answers are to these questions, if the account that I have proposed is correct, then the activity of the Kantian rational agent is much more unified than might at first appear, especially given Kant's tendency to focus on the differences between the practical and theoretical uses of our rational powers. The same basic set of conditions constituting the agent's control in exercising her power of spontaneity, and the connections between such control and the imputability of action, hold for the theoretical exercise of the intellect just as for the practical. For Kant, the exercise of one's rational capacities generally, and not just the rational exercise of one's will, thereby demonstrates a form of cognitive control, and thus requires the exercise of one's transcendental freedom.

Acknowledgments

This paper benefitted from a variety of input and advice. Special thanks to Derk Pereboom, from whose views over the years I've benefitted immensely. Thanks to

110. For discussion of these and related issues see Cohen (2014: chs. 4–5), Merritt (2018).

audience members at the 2018 Pacific APA, the University of Nebraska–Omaha, and Peking University for lively and helpful discussion. Thanks also to Yoon Choi, Andrew Roche, Karl Schafer, Eric Watkins, and two anonymous referees for this journal for extensive and helpful comments.

Translations & Abbreviations

Quotations from Kant's work are from the *Akademie Ausgabe*, with the first *Critique* cited by the standard A/B edition pagination, and the other works by volume and page. Where available, translations follow the Cambridge editions of the Works of Immanuel Kant, general editors Paul Guyer and Allen Wood. References to specific texts are abbreviated as follows:

An: *Anthropology from a Pragmatic Point of View*

C: *Correspondence*

CF: *Conflict of the Faculties*

CPJ: *Critique of the Power of Judgment*

CPR: *Critique of Pure Reason*

CPrR: *Critique of Practical Reason*

G: *Groundwork for the Metaphysics of Morals*

JL: *Jäsche Logic*

MM: *Metaphysics of Morals*

NE: *A New Elucidation of the First Principles of Metaphysical Cognition*

Pr: *Prolegomena to Any Future Metaphysics*

Refl: *Notes and Fragments*

Rel: *Religion within the Boundaries of Reason Alone*

References

- Adams, Robert Merrihew (1994). *Leibniz: Determinist, Theist, Idealist*. Oxford University Press.
- Allais, Lucy (2015). *Manifest Reality: Kant's Idealism and His Realism*. Oxford University Press.
- Allison, Henry E. (1973). *The Kant-Eberhard Controversy*. The Johns Hopkins University Press.
- Allison, Henry E. (1990). *Kant's Theory of Freedom*. Cambridge University Press.
- Allison, Henry E. (1996a). Autonomy and Spontaneity in Kant's Conception of the Self. In *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (129–142). Cambridge University Press.
- Allison, Henry E. (1996b). *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge University Press.

- Allison, Henry E. (1996c). Kant's Refutation of Materialism. In *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (92–106). Cambridge University Press.
- Allison, Henry E. (2004). *Kant's Transcendental Idealism: Revised and Enlarged*. Yale University Press.
- Allison, Henry E. (2006). Kant on Freedom of the Will. In Paul Guyer (Ed.), *The Cambridge Companion to Kant and Modern Philosophy* (381–415). Cambridge University Press.
- Alston, William P. (1988). The Deontological Conception of Epistemic Justification. *Philosophical Perspectives*, 2, 257–299.
- Ameriks, Karl (1992). Kant and Hegel on Freedom: Two New Interpretations. *Inquiry*, 35(2), 219–232.
- Ameriks, Karl (2000a). *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy*. Cambridge University Press.
- Ameriks, Karl (2000b). *Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason*. Oxford University Press.
- Ameriks, Karl (2003a). *Interpreting Kant's Critiques*. Oxford University Press.
- Ameriks, Karl (2003b). Kant's Deduction of Freedom and Morality. In *Interpreting Kant's Critiques* (161–192). Oxford University Press.
- Ameriks, Karl (2012). *Kant's Elliptical Path*. Clarendon Press.
- Bayne, Tim (2004). Self-Consciousness and the Unity of Consciousness. *The Monist*, 87(2), 219–236.
- Beck, Lewis White (1987). Five Concepts of Freedom in Kant. In Jan T. J. Szrednicki (Ed.), *Stephan Körner – Philosophical Analysis and Reconstruction* (35–51). Springer.
- Bermúdez, José Luis (2005). *Philosophy of Psychology: A Contemporary Introduction*. Routledge.
- Bilgrami, Akeel (2006). *Self-Knowledge and Resentment*. Harvard University Press.
- Bird, Graham (2006). *The Revolutionary Kant: A Commentary on the Critique of Pure Reason*. Open Court.
- Boghossian, Paul (2014). What Is Inference. *Philosophical Studies*, 169(1), 1–18.
- Boyle, Matthew (2009). Active Belief. *Canadian Journal of Philosophy*, 39(sup1), 119–147.
- Brewer, Kimberly and Eric Watkins (2012). A Difficulty Still Awaits: Kant, Spinoza, and the Threat of Theological Determinism. *Kant-Studien*, 103(2), 163–187.
- Broome, John (2013). *Rationality through Reasoning*. Wiley Blackwell.
- Broome, John (2014). Comments on Boghossian. *Philosophical Studies*, 169(1), 19–25.
- Butts, Robert E. (1997). Kant's Dialectic and the Logic of Illusion. In Patricia A. Easton (Ed.), *Logic and the Workings of the Mind: The Logic of Ideas and Faculty Psychology in Early Modern Philosophy* (307–317). Ridgeview.

- Byrne, Alex (2018). *Transparency and Self-Knowledge*. Oxford University Press.
- Carnois, Bernard (1987). *The Coherence of Kant's Doctrine of Freedom*. University of Chicago Press.
- Carruthers, Peter (2011). *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford University Press.
- Chignell, Andrew (2007a). Belief in Kant. *The Philosophical Review*, 116(3), 323–360.
- Chignell, Andrew (2007b). Kant's Concepts of Justification. *Noûs*, 41(1), 33–63.
- Chignell, Andrew (2014). Modal Motivations for Noumenal Ignorance: Knowledge, Cognition, and Coherence. *Kant-Studien*, 105(4), 573–597.
- Choi, Yoon (2019). Spontaneity and Self-Consciousness in the Groundwork and the B-Critique. *Canadian Journal of Philosophy*, 49(7), 936–955.
- Cohen, Alix (2013). Kant on Doxastic Voluntarism and Its Implications for Epistemic Responsibility. *Kant Yearbook*, 5(1), 33–50.
- Cohen, Alix (2014). Kant on the Ethics of Belief. *Proceedings of the Aristotelian Society*, 114(3 pt 3), 317–334.
- Coliva, Annalisa (2002). Thought Insertion and Immunity to Error Through Misidentification. *Philosophy, Psychiatry, & Psychology*, 9(1), 27–34.
- Conant, James (1992). The Search for Logically Alien Thought: Descartes, Kant, Frege, and the Tractatus. *Philosophical Topics*, 20(1), 115–180.
- Dennett, Daniel C. (1969). *Content and Consciousness*. Routledge & Kegan Paul.
- Duncan, Matt (2019). The Self Shows Up in Experience. *Review of Philosophy and Psychology*, 10(2), 299–318.
- Dyck, Corey W. (2014). *Kant and Rational Psychology*. Oxford University Press.
- Engstrom, Stephen (1993). Allison on Rational Agency. *Inquiry*, 36(4), 405–418.
- Engstrom, Stephen (2006). Understanding and Sensibility. *Inquiry*, 49(1), 2–25.
- Engstrom, Stephen (2009). *The Form of Practical Knowledge*. Harvard University Press.
- Frierson, Patrick R. (2014). *Kant's Empirical Psychology*. Cambridge University Press.
- Gallagher, Shaun (2000). Self-Reference and Schizophrenia: A Cognitive Model of Immunity to Error Through Misidentification. In Dan Zahavi (Ed.), *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-Experience* (203–241). Benjamins.
- Grier, Michelle (2001). *Kant's Doctrine of Transcendental Illusion*. Cambridge University Press.
- Grüne, Stefanie (2009). *Blinde Anschauung*. Vittorio Klostermann.
- Guyer, Paul (1987). *Kant and the Claims of Knowledge*. Cambridge University Press.
- Guyer, Paul (2009). Problems with Freedom: Kant's Argument in Groundwork III and Its Subsequent Emendations. In Jens Timmermann (Ed.), *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide* (176–202). Cambridge University Press.

- Hartz, Glenn A. and Jan A. Cover (1988). Space and Time in the Leibnizian Metaphysic. *Noûs*, 22(4), 493–519.
- Hasker, William (1973). The Transcendental Refutation of Determinism. *The Southern Journal of Philosophy*, 11(3), 175–183.
- Heis, Jeremy (2013). The Priority Principle from Kant to Frege. *Noûs*, 48(2), 268–297.
- Henrich, Dieter (1960). Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft. In Dieter Henrich, W Schulz, and K-H Volkmann-Schluck (Eds.), *Die Gegenwart der Griechen im neueren Denken: Festschrift für Hans-Georg Gadamer zum 60. Geburtstag* (77–115). Mohr.
- Henrich, Dieter (1975). Die Deduktion des Sittengesetzes: Über die Gründe der Dunkelheit des letzten Abschnittes von Kants, Grundlegung zur Metaphysik der Sitten. In Alexander Schwan (Ed.), *Denken im Schatten des Nihilismus. Festschrift für Wilhelm Weischedel* (55–112). Wissenschaftliche Buchgesellschaft.
- Hogan, Desmond (2009). How to Know Unknowable Things in Themselves. *Noûs*, 43(1), 49–63.
- Hogan, Desmond (2014). Kant on Foreknowledge of Contingent Truths. *Res Philosophica*, 91(1), 47–70.
- Hudson, Hud (1991). Wille, Willkür, and the Imputability of Immoral Actions. *Kant-Studien*, 82(2), 179–196.
- Jauernig, Anja (2010). Disentangling Leibniz's Views on Relations and Extrinsic Denominations. *Journal of the History of Philosophy*, 48(2), 171–205.
- Johnson, Robert N. (2009). The Moral Law as Causal Law. In Jens Timmermann (Ed.), *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide* (82–101). Cambridge University Press.
- Jorati, Julia (2015a). Leibniz on Causation – Part 1. *Philosophy Compass*, 10(6), 389–397.
- Jorati, Julia (2015b). Leibniz on Causation – Part 2. *Philosophy Compass*, 10(6), 398–405.
- Jorati, Julia (2017). *Leibniz on Causation and Agency*. Cambridge University Press.
- Jordan, James N. (1969). Determinism's Dilemma. *The Review of Metaphysics*, 23(1), 48–66.
- Kain, Patrick (2010). Practical Cognition, Intuition, and the Fact of Reason. In Benjamin Lipscomb and James Krueger (Eds.), *Kant's Moral Metaphysics: God, Freedom, and Immortality* (211–230). De Gruyter.
- Kain, Patrick (in press). The Development of Kant's Conception of Divine Freedom. In Brandon C. Look (Ed.), *Leibniz and Kant*. Oxford University Press.
- Keller, Pierre (1998). *Kant and the Demands of Self-Consciousness*. Cambridge University Press.
- Kitcher, Patricia (1990). *Kant's Transcendental Psychology*. Oxford University Press.
- Kitcher, Patricia (2011). *Kant's Thinker*. Oxford University Press.

- Kitcher, Patricia (2014). Replies. *Kantian Review*, 19(1), 149–159.
- Klemme, Heiner F. (1999). Kants Wende zum Ich. *Zeitschrift für philosophische Forschung*, 53(4), 507–529.
- Kohl, Markus (2015a). Kant and ‘Ought Implies Can’. *The Philosophical Quarterly*, 65(261), 690–710.
- Kohl, Markus (2015b). Kant on Freedom of Empirical Thought. *Journal of the History of Philosophy*, 53(2), 301–326.
- Kohl, Markus (in press). Kant on Cognizing Oneself as a Spontaneous Cognizer. *Canadian Journal of Philosophy*.
- Korsgaard, Christine M. (1986). Skepticism about Practical Reason. *The Journal of Philosophy*, 83(1), 5.
- Korsgaard, Christine M. (1989). Morality as Freedom. In Yimiyahu Yovel (Ed.), *Kant’s Practical Philosophy Reconsidered* (23–48). Springer.
- Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Korsgaard, Christine M. (1997). The Normativity of Instrumental Reason. In Garrett Cullity and Berys Nigel Gaut (Eds.), *Ethics and Practical Reason* (215–254). Clarendon Press.
- Kosch, Michelle (2006). *Freedom and Reason in Kant, Schelling and Kierkegaard*. Oxford University Press.
- Kreines, James (2008). Kant on the Laws of Nature: Laws, Necessitation, and the Limitation of Our Knowledge. *European Journal of Philosophy*, 17(4), 527–558.
- Kreines, James (2017). Kant on the Laws of Nature: Restrictive Inflationism and Its Philosophical Advantages. *The Monist*, 100(3), 326–341.
- Kriegel, Uriah (2004). Consciousness and Self-Consciousness. *The Monist*, 87(2), 182–205.
- Langton, Rae (1998). *Kantian Humility: Our Ignorance of Things in Themselves*. Oxford University Press.
- Leech, Jessica (2015). Logic and the Laws of Thought. *Philosophers’ Imprint*, 15(12), 1–27.
- Longuenesse, Béatrice (1998). *Kant and the Capacity to Judge*. Princeton University Press.
- Longuenesse, Béatrice (2005). *Kant on the Human Standpoint*. Cambridge University Press.
- Longuenesse, Béatrice (2017). *I, Me, Mine: Back to Kant, and Back Again*. Oxford University Press.
- Lu-Adler, Huaping (2017). Kant and the Normativity of Logic. *European Journal of Philosophy*, 25(2), 207–230.
- MacFarlane, John (2002). Frege, Kant, and the Logic in Logicism. *Philosophical Review*, 111(1), 25–65.
- McCarty, Richard (2009). *Kant’s Theory of Action*. Oxford University Press.

- McCormick, Miriam Schleifer (2017). Responding to Skepticism About Doxastic Agency. *Erkenntnis*, 83(4), 627–645.
- McHugh, Conor (2012). Epistemic Deontology and Voluntariness. *Erkenntnis*, 77(1), 65–94.
- McHugh, Conor and Jonathan Way (2016). Against the Taking Condition. *Philosophical Issues*, 26(1), 314–331.
- McKenna, Michael (2001). Source Incompatibilism, Ultimacy, and the Transfer of Non-Responsibility. *American Philosophical Quarterly*, 38(1), 37–51.
- McLear, Colin (2019a). Apperception & the Substantial Subject. Manuscript in preparation.
- McLear, Colin (2019b). Kant on the Reflective Nature of Reasoning. Manuscript in preparation.
- McLear, Colin (2019c). The Unity of Kant’s Conception of Reflection. Manuscript in preparation.
- McLear, Colin (2016). Kant on Perceptual Content. *Mind*, 125(497), 95–144.
- McLear, Colin (2019). The Mind’s “I”. *European Journal of Philosophy*, 27(1), 255–265.
- McLear, Colin (in press). Animals & Objectivity. In Lucy Allais and John Callanan (Eds.), *Kant on Animals*. Oxford University Press.
- McLear, Colin and Derk Pereboom (in press). Kant on Transcendental Freedom, Priority Monism, & the Structure of Intuition. In Evan Tiffany and Dai Heide (Eds.), *The Idea of Freedom: New Essays on the Kantian Theory of Freedom*. Oxford University Press.
- McRae, Robert (1976). *Leibniz: Perception, Apperception, and Thought*. University of Toronto Press.
- Meier, Georg Friedrich (2016). *Excerpt from the Doctrine of Reason*. Bloomsbury.
- Merritt, Melissa McBay (2018). *Kant on Reflection and Virtue*. Cambridge University Press.
- Merritt, Melissa McBay and Markos Valaris (2017). Attention and Synthesis in Kant’s Conception of Experience. *The Philosophical Quarterly*, 67(268), 571–592.
- Messina, James (2017). Kant’s Necessitation Account of Laws and the Nature of Natures. In Michela Massimi and Angela Breitenbach (Eds.), *Kant and the Laws of Nature* (131–149). Cambridge University Press.
- Mischel, Walter, Ebbe B. Ebbesen, and Antonette Raskoff Zeiss (1972). Cognitive and Attentional Mechanisms in Delay of Gratification. *Journal of Personality and Social Psychology*, 21(2), 204–218.
- Mohr, Georg (1991). *Das sinnliche Ich: innerer Sinn und Bewusstsein bei Kant*. Königshausen & Neumann.
- Mohr, Georg (1995). Freedom and the Self: From Introspection to Intersubjectivity: Wolff, Kant, and Fichte. In Karl Ameriks and Dieter Sturma (Eds.), *The Modern Subject: Conceptions of the Self in Classical German Philosophy* (31–45). SUNY

- Press.
- Neta, Ram (2019). The Basing Relation. *The Philosophical Review*, 128(2), 179–217.
- Nunez, Tyke (2018). Logical Mistakes, Logical Aliens, and the Laws of Kant’s Pure General Logic. *Mind*, 128(512), 1149–1180.
- Pasternack, Lawrence (2011). The Development and Scope of Kantian Belief: The Highest Good, the Practical Postulates and the Fact of Reason. *Kant-Studien*, 102(3), 290–315.
- Pasternack, Lawrence (2014). Kant on Opinion: Assent, Hypothesis, and the Norms of General Applied Logic. *Kant-Studien*, 105(1), 41–82.
- Paton, H. J. (1947). *The Categorical Imperative: A Study in Kant’s Moral Philosophy*. Hutchinson.
- Pereboom, Derk (1995). Self-Understanding in Kant’s Transcendental Deduction. *Synthese*, 103(1), 1–42.
- Pereboom, Derk (2001). *Living without Free Will*. Cambridge University Press.
- Pereboom, Derk (2006a). Kant on Transcendental Freedom. *Philosophy and Phenomenological Research*, 73(3), 537–567.
- Pereboom, Derk (2006b). Kant’s Metaphysical and Transcendental Deductions. In Graham Bird (Ed.), *A Companion to Kant* (154–168). Blackwell.
- Pereboom, Derk (2014). *Free Will, Agency, and Meaning in Life*. Oxford University Press.
- Pippin, Robert B. (1987). Kant on the Spontaneity of Mind. *Canadian Journal of Philosophy*, 17(2), 449–475.
- Pollok, Konstantin (2017). *Kant’s Theory of Normativity: Exploring the Space of Reason*. Cambridge University Press.
- Prauss, Gerold (1983). *Kant über Freiheit als Autonomie*. Vittorio Klostermann.
- Radden, Jennifer (1998). Pathologically Divided Minds, Synchronic Unity and Models of Self. *Journal of Consciousness Studies*, 5(5–6), 658–672.
- Rödl, Sebastian (2007). *Self-Consciousness*. Harvard University Press.
- Rohlf, Michael (2010). The Ideas of Pure Reason. In Paul Guyer (Ed.), *The Cambridge Companion to Kant’s Critique of Pure Reason* (190–209). Cambridge University Press.
- Rosefeldt, Tobias (2000). *Das logische Ich: Kant über den Gehalt des Begriffes von sich selbst*. PHILO.
- Rosenkoetter, Timothy (2019). The Canon Problem and the Explanatory Priority of Capacities. *Logical Analysis and History of Philosophy: From Leibniz to Kant*, 21, 216–234.
- Rutherford, Donald (2013). Laws and Powers in Leibniz. In Eric Watkins (Ed.), *The Divine Order, the Human Order, and the Order of Nature: Historical Perspectives* (149–174). Oxford University Press.
- Ryle, Gilbert (2009). *The Concept of Mind*. Routledge.
- Santi, Marco (2018). Kant on Epistemic Belief. Manuscript in preparation.

- Sartorio, Carolina (2013). Making a Difference in a Deterministic World. *The Philosophical Review*, 122(2), 189–214.
- Sartorio, Carolina (2016). *Causation and Free Will*. Oxford University Press.
- Schafer, Karl (in press-a). Kant's Conception of Cognition and our Knowledge of Things-in-Themselves. In Karl Schafer and Nicholas Stang (Eds.), *The Sensible and Intelligible Worlds: New Essays on Kant's Metaphysics and Epistemology*. Oxford University Press.
- Schafer, Karl (in press-b). Practical Cognition and Knowledge of Things-in-Themselves. In Dai Heide and Evan Tiffany (Eds.), *The Idea of Freedom: New Essays on the Kantian Theory of Freedom*. Oxford University Press.
- Schönecker, Dieter (1999). *Kant–Grundlegung III: die Deduktion des kategorischen Imperativs*. Alber.
- Schönecker, Dieter (2005). *Kants Begriffstranzendentaler und praktischer Freiheit: Eine entwicklungsgeschichtliche Studie*. De Gruyter.
- Sellars, Wilfrid (1971). "... This I or He or It (The Thing) Which Thinks ...". *Proceedings and Addresses of the American Philosophical Association*, 44, 5–31.
- Siegel, Susanna (2010). The Contents of Perception. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/archives/fall2010/entries/perception-contents/>
- Siegel, Susanna (2017a). Inference without Reckoning. In *The Rationality of Perception* (77–106). Oxford University Press.
- Siegel, Susanna (2017b). *The Rationality of Perception*. Oxford University Press.
- Smit, Houston (2000). Kant on Marks and the Immediacy of Intuition. *The Philosophical Review*, 109(2), 235–266.
- Smit, Houston (2009). Kant on Apriority and the Spontaneity of Cognition. In Samuel Newlands and Larry M. Jorgensen (Eds.), *Metaphysics and the Good: Themes from the Philosophy of Robert Adams (188–251)*. Oxford University Press.
- Smit, Houston (2019). Kant's "I Think" and the Agential Approach to Self-Knowledge. *Canadian Journal of Philosophy*, 49(7), 980–1011.
- Smith, Angela M. (2005). Responsibility for Attitudes: Activity and Passivity in Mental Life. *Ethics*, 115(2), 236–271.
- Stang, Nicholas F. (2016a). *Kant's Modal Metaphysics*. Oxford University Press.
- Stang, Nicholas F. (2016b). Kant's Transcendental Idealism. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/kant-transcendental-idealism/>
- Stang, Nicholas F. (2019). A Guide to Ground in Kant's Lectures on Metaphysics. In Courtney D. Fugate (Ed.), *Kant's Lectures on Metaphysics: A Critical Guide* (74–101). Cambridge University Press.
- Stephens, G. Lynn and George Graham (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. MIT Press.
- Stern, Robert (2004). Does 'Ought' Imply 'Can'? And Did Kant Think It Does?

- Utilitas*, 16(1), 42–61.
- Stevenson, Leslie (2003). Opinion, Belief or Faith, and Knowledge. *Kantian Review*, 7(1), 72–101.
- Stevenson, Leslie (2004). Freedom of Judgement in Descartes, Hume, Spinoza and Kant. *British Journal for the History of Philosophy*, 12(2), 223–246.
- Stich, Stephen P. (1978). Beliefs and Subdoxastic States. *Philosophy of Science*, 45(4), 499–518.
- Stratmann, Joe (2018). Kant, Grounding, and Things in Themselves. *Philosopher's Imprint*, 18(7), 1–21.
- Tenenbaum, Sergio (2012). The Idea of Freedom and Moral Cognition in Groundwork III. *Philosophy and Phenomenological Research*, 84(3), 555–589.
- Thompson, Manley (1981). On A Priori Truth. *The Journal of Philosophy*, 78(8), 458–482.
- Thorpe, Lucas (2011). Kant on the Relationship Between Autonomy and Community. In Charlton Payne and Lucas Thorpe (Eds.) *Kant and The Concept of Community* (63–87). University of Rochester Press.
- Timmermann, Jens (2003). Sollen und Können. “Du kannst, denn du sollst” und “Sollen impliziert Können” im Vergleich. *Logical Analysis and History of Philosophy*, 6, 113–122.
- Tolley, Clinton (2006). Kant on the Nature of Logical Laws. *Philosophical Topics*, 34(1&2), 371–407.
- Tolley, Clinton (2008). Kant and the Normativity of Logic. *Recht und Frieden in der Philosophie Kants: Akten des X. Internationalen Kant-Kongresses*, 5, 215–227.
- Tolley, Clinton (2011). Kant on the Content of Cognition. *European Journal of Philosophy*, 22(2), 200–228.
- Van Cleve, James (1999). *Problems from Kant*. Oxford University Press.
- Van Inwagen, Peter (1983). *An Essay on Free Will*. Clarendon Press.
- Vance Buroker, Jill (2017). Kant on Judging and the Will. In Anil Gomes and Andrew Stephenson (Eds.), *Kant and the Philosophy of Mind: Perception, Reason, and the Self* (189–203). Oxford University Press.
- Vanzo, Alberto (2018). Leibniz on Innate Ideas and Kant on the Origin of the Categories. *Archiv für Geschichte der Philosophie*, 100(1), 19–45.
- Vetter, Barbara (2015). *Potentiality: From Dispositions to Modality*. Oxford University Press.
- Ware, Owen (2017). Kant's Deductions of Morality and Freedom. *Canadian Journal of Philosophy*, 47(1), 116–147.
- Watkins, Eric (2005). *Kant and the Metaphysics of Causality*. Cambridge University Press.
- Watkins, Eric (2010). Kant (on Action). In Timothy O'Connor and Constantine Sandis (Eds.), *A Companion to the Philosophy of Action* (521–527). Blackwell.
- Watkins, Eric (2019a). *Kant on Laws*. Oxford University Press.

- Watkins, Eric (2019b). What Real Progress Has Metaphysics Made since the Time of Kant? Kant and the Metaphysics of Grounding. *Synthese*. Advance online publication. <https://doi.org/10.1007/s11229-019-02180-2>
- Watkins, Eric and Marcus Willaschek (2017). Kant's Account of Cognition. *Journal of the History of Philosophy*, 55(1), 83–112.
- Watts, Tyler W., Greg J. Duncan, and Haonan Quan (2018). Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes. *Psychological Science*, 29(7), 1159–1177.
- Willaschek, Marcus (1992). *Praktische Vernunft: Handlungstheorie und Moralbegründung bei Kant*. J. B. Metzler.
- Willaschek, Marcus (2006). Die "Spontaneität des Erkenntnisses". Über die Abhängigkeit der "Transzendentalen Analytik" von der Auflösung der Dritten Antinomie. In Jiří Chotaš, Jindřich Karásek, and Jürgen Stolzenberg (Eds.), *Metaphysik und Kritik: Interpretationen zur "Transzendentalen Dialektik" der "Kritik der reinen Vernunft" (165–183)*. Königshausen & Neumann.
- Willaschek, Marcus and Eric Watkins (2017). Kant on Cognition and Knowledge. *Synthese*. Advance online publication. <https://doi.org/10.1007/s11229-017-1624-4>
- Williams, Bernard (1995). Internal Reasons and the Obscurity of Blame. In *Making Sense of Humanity: And Other Philosophical Papers 1982–1993* (35–45). Cambridge University Press.
- Wolt, Daniel (2018). Kant on Free Will and Theoretical Rationality. *Ideas y Valores*, 67(166), 181–198.
- Wood, Allen W. (1970). *Kant's Moral Religion*. Cornell University Press.
- Wood, Allen W. (1984). Kant's Compatibilism. In Allen W. Wood (Ed.), *Self and Nature in Kant's Philosophy* (73–101). Cornell University Press.
- Wood, Allen W. (1999). *Kant's Ethical Thought*. Cambridge University Press.
- Wood, Allen W. (2008). *Kantian Ethics*. Cambridge University Press.
- Wright, Crispin (2014). Comment on Paul Boghossian, "What Is Inference". *Philosophical Studies*, 169(1), 27–37.
- Wuerth, Julian (2014). *Kant on Mind, Action, and Ethics*. Oxford University Press.