TRANSCENDENTAL IDEALISM AS A FRAMEWORK FOR AGENT-CAUSAL LIBERTARIANISM

A Dissertation Submitted

to the

Temple University Graduate Board

In Partial Fulfillment

of Requirements

for the Degree DOCTOR OF PHILOSOPHY

---

by Daniel Paul Dal Monte

December 2020

Examining Committee Members:

Dr. Lara Ostaric, Advisory Chair, Philosophy

Dr. Espen Hammer, Philosophy

Dr. Eugene Chislenko, Philosophy

Dr. Eric Watkins, External Member, San Diego, University of California-San Diego

Dr. James Hebbeler, External Member, Philadelphia, St. Joseph's University

ABSTRACT

In this dissertation, I occupy two realms of philosophy that have not been commonly associated. On the one hand, I enter into debates about the proper interpretation of Kant, specifically having to do with the very fractured debate on the nature and applications of transcendental idealism. I adjudicate on the matters of the relationship between appearances and things in themselves, i.e. whether it is epistemological or ontological, the way in which TI resolves the antinomial conflict of reason as it thinks the unconditioned in its exploration of cosmological questions, and the way Kant applies TI to articulate the intelligible and empirical characters in his metaphysics of agency. In addition to this historical research, I also turn to contemporary formulations of libertarian freedom. Libertarianism in free will debates is the view that free will is incompatible with determinism (i.e. incompatibilism), and free will exists. Libertarianism is a competitor to compatibilist views that claim that free will is compatible with determinism, i.e. the view that there is a unique outcome given a past and laws of nature characterizing the past.

Within libertarianism, there are important differences in terms of the metaphysics of free will. Most contemporary libertarians opt for a reductionist metaphysics, in which causation consists in relationships between events and does not involve underlying grounds or substances. Both event-causal and non-causal libertarianism accordingly ground their views of freedom on the interplay of psychological events conceived of as states of affairs at instants in time. Event-causal accounts locate free will in indeterministic causal series, involve conflicting sets of motivations that resolve

themselves probabilistically into a certain kind of action. Non-causal accounts do not attach free will to causality at all, instead associating it with a spontaneously occurring event.

ECL and NCL struggle with establishing how the agent actually settles her action. If the action is merely the indeterministic byproduct of a set of psychological processes or process, then what ultimately occurs is not up to the agent but a product of chance. ECL and NCL nevertheless object that even an action settled by chance is done consciously and according to reasons. But these criteria are aligned with compatibilist criteria for free will. Compatibilists deny to the agent the unconditioned power to choose, which is independent of any prior determination but also not subject to chance. They point out, though, that the action is externally unconstrained, or that, if the reasons had been different, the agent would have acted differently. Since contemporary libertarianism deprives the agent of control, and creates a kind of pseudo-agent that acts ultimately according to chance, I explore other metaphysical frameworks for free will.

Agent-causal libertarianism involves the agent directly causing her action as a substance. It is not some state of affairs that causes the action—a desire or belief characterizing the agent's psychology at a certain time—but the agent herself. Agent-causation promises to resolve the problem of control associated with event-causation. The agent-caused action is neither produced deterministically from a prior event, nor is it an indeterministic fallout from probabilistic causation. Instead, it is caused by an agent-substance able to act independently of events.

Timothy O'Connor is a well-known and articulate defender of agent-causation, but he also subscribes to the naturalistic framework popular in contemporary

metaphysics. Even though he accepts the reality of emergent properties, the agent-cause, which has a special capacity for self-determination, is supposed to be causally united to a microphysical level where there is only passive event-causation.

In this dissertation, I seek to frame agent-causation in terms of transcendental idealism. Rather than establishing the level of event-causation as metaphysically fundamental, I explore an idealistic metaphysics in which the empirical world in the spatiotemporal framework of human experience is not an absolute measure of reality. The human person is a hybrid creature, spanning two domains. On the one hand, the person exists in the empirical order in space and time. It is characterized by events subject to a causal law, by which they are accounted for in terms of prior events. On the other hand, there is a deeper level to the person, not encompassed within the limited structures of human experience. On this deep intelligible level, the human person is able to serve as the unconditioned ground of its empirical character.

DEDICATION

For Those Who Seek to Know and Live by the Truth

ACKNOWLEDGMENTS

This dissertation could not have come together were it not for the assistance of several people. Dr. Lara Ostaric has been an excellent advisor. She gave each chapter a very careful examination and provided insightful commentary, even as initial drafts were rough and I was having difficulty formulating my ideas. I also would like to acknowledge Dr. Ostaric for pointing out the importance of metaphysics in Kant, in both her seminars and published work, as many contemporary commentators seek to annihilate a metaphysical distinction between the world of appearances and things in themselves, and any suggestion that the ideas of reason could have an objective metaphysical realization. Though I have never met Dr. Karl Ameriks, I would also like to acknowledge his work, since he was Lara's dissertation advisor and also advised two other members of my committee, Dr. Eric Watkins and Dr. James Hebbeler. I have read several of Dr. Ameriks's books and found him a crucial support in combating the short argument to idealism that has been part of Kant commentary since Reinhold. The short argument reaches things in themselves through the mere epistemic fact of representation, which entails the unrepresentable. The short argument does not have to ascribe a metaphysical limitedness to the spatiotemporal framework of human experience.

I would also like to thank Dr. Espen Hammer. Professor Hammer allowed me to attend his seminar on Kant in the spring semester of 2020. I benefited greatly from this seminar, especially the discussion of the Transcendental Deduction. These discussions became the inspiration of my critique of the Neo-Humean empiricism of contemporary libertarianism in Chapter 4 of this dissertation, in which I appeal to a regressive strategy

to argue that empirical knowledge cannot be grounded in sensations alone but requires an a priori metaphysical structure.

I also need to mention Dr. Eric Watkins, who provided me invaluable study opportunities with the University of California-San Diego faculty. Through Professor Watkins's interest in my development, I was able to attend a seminar in Mainz, Germany, and also one in San Diego with Dr. Lucy Allais. I have also benefited greatly from Professor Watkins's 2005 work, *Kant and the Metaphysics of Causality*.

A NOTE ON SOURCES KANT

Apart from the Kritik der reinen Vernunft [Critique of Pure Reason], all references to Kant's writings are to the appropriate volume (in upper case roman numerals) and page number of Kants gesammelte Schriften (AA), edited by the Deutsche Akademie der Wissenschaften zu Berlin (formerly the Königlichen Preussischen Akademie der Wissenschaften), 29 vols. (Berlin: de Gruyter, 1900-). References to the Kritik der reinen Vernunft are to the standard A and B pagination of the first and the second editions (1781 and 1787 respectively). Citations to Reflexionen are made by reference to the reflection number, followed by the relevant volume and page number in AA.

Below is the list of abbreviations of Kant's works together with the information on the English translations I have consulted.

Diss De Mundi Sensibilis Atque Intelligibis Forma Et Principiis (KGS 2). On the Form and Principles of the Sensible and Intelligible World (The Inaugural Dissertation), trans. David Walford and Ralf Meerbote, Cambridge: Cambridge University Press, 2002.

GMS Grundlegung zur Metaphysik der Sitten. 1st ed., 1785; 2nd ed., 1786. ["Groundwork of the Metaphysics of Morals," in Immanuel Kant Practical Philosophy, ed. and trans. Mary J. Gregor. Cambridge University Press, Cambridge, 1996.]

KdpV Kritik der praktischen Vernunft. 1788. ["Critique of Practical Reason," in Immanuel Kant Practical Philosophy, ed. and trans. Mary J. Gregor. Cambridge University Press, Cambridge, 1996.]

KdrV Kritik der reinen Vernunft. 1st ed. [A], 1781; 2nd ed. [B], 1787. [Critique of Pure Reason, ed. and trans. Allen W. Wood and Paul Guyer. Cambridge University Press,

Cambridge, 1998.] **KdU** Kritik der Urteilskraft. 1st ed., 1790; 2nd ed., 1793; 3rd ed., 1799. [Critique of Judgment, ed. Paul Guyer, trans. Paul Guyer and Eric Matthews. Cambridge University Press, Cambridge, 2000.]

Lk Gedanken von der wahren Schätzung der lebendigen Kräfte. 1747. [Thoughts on the True Estimation of Living Forces, ed. Eric Watkins and trans. Jeffrey Edwards and Martin Schönfeld. Cambridge University Press, Cambridge, 2012.}

MadN Metaphysische Anfangsgrunde der Naturwissenschaft. 1783. [Metaphysical Foundations of Natural Science, ed. and trans. Michael Friedman, Cambridge University Press, Cambridge, 2004.]

MdS Die Metaphysik der Sitten. 1st ed., 1797; 2nd ed., 1798. ["The Metaphysics of Morals," in Immanuel Kant Practical Philosophy, ed. and trans. Mary J. Gregor. Cambridge University Press, Cambridge, 1996.]

ND New Elucidation, *Principiorum Primorum Cognitionis Metaphysicae Nova Dilucidatio*. ["New Elucidation," in Theoretical Philosophy, ed. and trans. by David Walford and Ralf Meerbote. Cambridge University Press, Cambridge, 2002.]

P Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können. 1783. [Prolegomena to Any Future Metaphysics That Will Be Able to Present Itself as Science, ed. Günter Zöller, trans. Peter G. Lucas and Günter Zöller. Oxford University Press, Oxford, 2004.]

R Die Religion innerhalb der Grenzen der blossen Vernunft. 1st ed., 1793; 2nd ed., 1794. ["Religion within the Boundaries of Mere Reason," in Kant Religion within the

Boundaries of Mere Reason and Other Writings, ed. Allen W. Wood and George di Giovanni. Cambridge University Press, Cambridge, 1998.]

WhdO Wass heißt sich im Denken orientiren? 1786. ["What Does It Mean to Orient Oneself in Thinking?" in Religion within the Boundaries of Mere Reason and Other Writings, ed. Allen W. Wood and George di Giovanni. Cambridge University Press, Cambridge, 1998.]

TABLE OF CONTENTS

Page

PREFACE

AGENT-CAUSATION WITHIN THE FRAMEWORK OF TRANSCENDENTAL IDEALISM

Agent-causation is a theory of freedom that provides a potentially attractive solution to problems associated with other metaphysics of agency. Agent-causation involves direct causation by an agent of its action. The agent, in agent-causal accounts, is able to directly the settle what its actions will be, independently of any events that might characterize it—e.g. desires, beliefs, and other motivating factors that might arise at certain times.[1] Agent-causal theories provide a corrective to the so-called luck objection to event-causal theories, which account for the origination of action through agent-involving events.[2] The problem of event-causal accounts, though, is that, if an action is driven only be agent-involving events, then it appears as though the action is not under the control of the agent. Whether the series of events proceed deterministically or indeterministically, it appears as though the actions they produce are based entirely on the luck of the circumstances in which the agent happens to find herself.[3]

Agent-causation might seem an obvious solution to the lack of agential control in the event-causal accounts, were it not for problems with agent-causation that make it, to some, even more repellent. Agent-causal accounts have been largely abandoned in contemporary free will theory, for two reasons identified by a defender of agent-

[1] Defenders of agent-causation in the past couple of centuries include Immanuel Kant, Thomas Reid, Roderick Chisholm, Richard Taylor, Timothy O'Connor, Randolph Clarke, and Meghan Griffith. This is mentioned in Derk Pereboom and Michael McKenna (2016): 243.
[2] Contemporary defenders of event-causal accounts are Mark Balaguer and Robert Kane. Ibid., 233.
[3] Luck objections to event-causal accounts can be found in David Hume, as well as contemporary philosophers like Peter van Inwagen. Ibid., 236.

causation, Randolph Clarke.[4] First, agent-causal accounts fail to articulate how rational free action is possible, because they fail to articulate a metaphysics of how an agent, her reasons, and her action are interrelated. If an agent-cause is supposed to directly select an action independently of events, then it seems as though the selection would have to be arbitrary. Certain motivational sets could not drive the selection, because these are agent-involving events the agent-causal theorist abjures as the cause of action.

The second reason for the abandonment of agent-causal accounts, according to Clarke, is that they do not provide an intelligible account of that in which direct causation by an agent is supposed to consist. If agent-causation does not occur through some event, then it seems as though it has no mechanism by which it exercises control. Empiricists in contemporary philosophy would therefore reject the opaque agent-cause as a non-experiential fabrication.

My thesis in this dissertation is that the Kantian doctrine of transcendental idealism can provide a compelling framework that can resolve these two, as well as other, difficulties with agent-causal theories. In the first section of this introduction, I will identify two forms of autonomy in Kant, theoretical and practical, and how they are related. It is the proper integration of theoretical and practical autonomy that can provide a sound foundation of a theory of agent-causation. In the second section, I will explain how theoretical autonomy and practical autonomy, in contemporary Kant scholarship, have been juxtaposed in an oppositional relationship. In the third section, I will explain the disastrous consequences of setting theoretical autonomy against practical autonomy

[4] See Randolph Clarke (1993): 191.

for the metaphysics of agency. In the fourth and final section, I provide an overview of the chapter layout in this dissertation.

Theoretical and Practical Autonomy in Kant and Their Relationship

*Theoretical Autonomy***.** In Kant, there is an emphasis on the spontaneity of one's theoretical faculties, which consists in the fact that one's understanding takes an active synthesizing role in one's experience. The constructivist role of the understanding in synthesizing intuitions through pure concepts is a rejection of the so-called Myth of the Given. The Myth of the Given, broadly speaking, is that cognitive states arise directly from sensory impressions. Those who debunk the Myth of the Given recognize that cognitive states, in which we formulate propositions that can be true or false, are categorically different from mere sensory impressions, which cannot be true or false.[5]

Kant debunks the Myth of the Given in his position that intuitions without concepts are blind.[6] In order to enter into propositional cognitive states and justificatory relationships, intuitions must be taken up by the pure concepts of the understanding. The understanding, which structures intuitions according to concepts, and sensibility, which receives intuitions, differ in terms of their receptivity and spontaneity. While sensibility is receptive, since it only involves a passive relationship with objects, the understanding is

[5] For an account of the myth of the given from the perspectives of both Sellars and McDowell, see Eric Watkins (2008): 516.
[6] The famous formulation of this position comes in *KdrV* A51/B75.

spontaneous, insofar as it creates full-fledged cognitive states by bringing to bear concepts on sensible intuitions.[7]

The fact that cognitive content for Kant comes through a union of pure concepts and sensible intuition creates a limitation for human cognition that has important implications for practical autonomy/spontaneity. For Kant, we cannot pursue knowledge of objects in the world through our pure understanding, divorced from received sensible content. Accordingly, we cannot have cognition apart from the structure of sensibility, which, according to Kant's Transcendental Aesthetic, is necessarily spatiotemporal.[8] Moreover—though this step is controversial[9]—it appears as though Kant wants to establish that the spatiotemporality that is the universal feature of our sensibility is a merely contingent structure that is not an absolute structure of reality. Particularly in the First Antinomy, it appears that the spatiotemporal world cannot constitute an unconditioned totality with a definite magnitude and duration, either finite or infinite, because space and time do not exist apart from the empirical synthesis of the sensibility of the human subject.[10] It is because cognition takes place, then, only in the union of concepts and intuitions, which must be in space and time, and that space and time are only subject-relative and non-absolute structures, that cognition is only of appearances, not of things in themselves. We can abstract from the spatiotemporal conditions of appearances through thinking with the pure concepts of the understanding.[11]

---

[7] Ibid.
[8] Kant claims that the a priori forms of sensibility are space and time at *KdrV* A42/B59.
[9] Henry Allison, for instance, thinks that Kant is only concerned with establishing space and time as epistemic conditions, and is not interested in any alternative ontology about space and time. Henry Allison (2004): 98.
[10] The position that the spatiotemporal world is neither infinite nor finite comes at *KdrV* A520/B548.
[11] *KdrV* B166.

*Practical autonomy***.** Kant says in the second *Critique* that practical reason has cognition of itself as absolutely free. This absolute freedom is a kind of freedom that is not subject to a causal law connecting it to previous events, and so it is not an object of possible experience. Theoretical cognition, therefore, cannot have cognition of absolute freedom. However, the fact that theoretical cognition pertains only to appearances in a contingent spatiotemporal domain allows us to be open to the possibility of a supersensible domain outside of the limits of possible experience.[12] The contingency of the spatiotemporal structure of sensibility means that practical reason can extend theoretical reason, in its cognition of itself as in possession of a supersensible faculty of absolute freedom.[13]

Practical and theoretical autonomy are therefore importantly related. We are able to integrate our practical cognition of ourselves as absolutely free only because we have placed the human subject at the center of experience. Making human subjectivity central involves making its synthesizing role take place in the contingent domain of the necessary structures of its sensibility.[14] The centrality of human subjectivity is the basis for the appearance/thing in itself distinction that prevents theoretical cognition's

[12] *KdpV* 5:5.

[13] The integration of our faculties in the form of a self-consciousness which applies its categories to the pure structure of sensibility, and absolute freedom, represents an integration of the interests of Kant to both advance knowledge to honor mankind, as well as to respect the rights of mankind. He derived the latter concern from Rousseau. See David Pacini's foreword of Henrich (2003): xxiii.

[14] Guyer discusses how Kant formulates a subjective metaphysics of experience, in which cognitive structures are conditional structures having to do only with experience and not existence as such, in Guyer (1987): 59.

experience of the continuity of the causal law[15] from crowding out the possibility of the absolute freedom of practical reason.

Though the integration of theoretical cognition and the absolute freedom of practical reason will be a project of this entire dissertation, I will note here how this integration can provide a way out of the seemingly intractable problems that have discredited agent-causal theories in contemporary philosophy.[16] The first problem I identified had to with integrating reasons into an agent-causal accounts that appeals to direct causation by an agent apart from an event. In a Kantian metaphysics, we do not entirely abjure event-causation, since it is still intact at the level of appearances. Kant states that we can predict with certainty the future actions of an individual given sufficient knowledge of empirical causes.[17] We can integrate absolute freedom, not as a direct cause of action, but as the transcendental cause of the event-causal order, since this order is not a metaphysical absolute.[18] The second problem I identified has to do with the opacity of the agent-cause. Though the agent-cause may not manifest itself in the form of an event, this does not mean it is opaque. The integration of a supersensible capacity for absolute freedom, and the contingency of the event-causal order, means that absolute freedom can act as a ground, rather than as another event in the temporal series.[19]

---

[15] Kant establishes that any event in succession in time must be subject to a causal law in the argument for the second Analogy of Experience, in *KdrV* B233.

[16] I think there is a larger context underlying the opposition between theoretical and practical spontaneity. This larger context has to do with the opposition between analytical and Continental philosophy. An integrated theoretical and practical reason could provide the systematic metaphysics characteristic of Continental philosophy, while at the same time retaining the empiricist perspective of analytic philosophy. For this delineation between analytic and continental, see Dieter Henrich (2003): 5.

[17] *KdrV* A550/B578.

[18] Kant uses the phrase, "transcendental cause," at *KdrV* A546/B574.

[19] Kant differentiates between origination in time and origination in reason alone in discussing how it is possible for an individual to have a subjective character that underlies all his individual empirical actions in time in *R* 6:39.

The Disunity of Practical and Theoretical Autonomy in Contemporary Kant Scholarship

In contemporary scholarship on Kant and Kantian themes, a Kantian modesty about the theoretical faculty, that makes the understanding a spontaneous faculty synthesizing intuitions instead of a receptive faculty tracking the nature of reality in itself, has turned into a cudgel against any kind of transcendent metaphysics associated with practical reason. For instance, theorists on Kantian ethics have divorced the self-legislation of the autonomous will from any kind of larger context of non-secular values. Though Kant states in *KdpV* that practical reason's access to the supersensible in the form of absolute freedom in turn allows it to continue to build a rationalist metaphysics involving both God and the immortal soul[20], contemporary philosophers have superimposed the limitations of theoretical cognition to sensible intuition on to the self-legislation of practical reason. Autonomous self-legislation is not to involve an appeal to transcendent metaphysics or non-secular values: we are to use entirely naturalistic procedures, such as reflective equilibrium, in which we find a coherence between individual decisions and our core values, or a purely human form of consensus-building.[21] In sum, the idea that the concepts of theoretical cognition are blind in the first *Critique* has the effect in contemporary scholarship of pushing out the metaphysics Kant associates with practical reason, i.e. an unconditional moral law that applies universally across any empirical situation (and so is more than just the consensus of a particular group, or the internal self-analysis of an individual), a form of absolute freedom that

[20] Absolute freedom's ability to lead us into a rationalist metaphysics involving supersensible objects makes it the keystone concept of the whole system of pure reason. *KdpV* 5:4.
[21] Karl Ameriks traces this tendency in Karl Ameriks (2000): 6.

undercuts the continuity of the conditioned-conditions relationship in experience, and religious convictions based on practical reason, such as a belief in God able to properly reward action in accordance with the moral law.[22]

We see the schizophrenia between the absolutization of the centrality of theoretical reason to experience, and the metaphysical commitments of practical reason, in the seminal work of Peter Strawson, *The Bounds of Sense*. Strawson identifies "two faces" of the first Critique. On the one hand, Strawson values the insights into what counts as an intelligible experience for us.[23] Our cognition has a certain limiting framework that forms the boundaries of what is intelligible for us. For instance, we cannot have experience outside of space and time. Nor is it possible, as Kant points out in his Transcendental Deduction, to have representations unless these are united in a single self-consciousness, i.e. a transcendental unity of the "I think" that accompanies every representation as an essential precondition.[24] On the other hand, Strawson rejects the other face he identifies in Kant, which is an idealistic metaphysics in which the limiting framework of our experience become tied in an ontological way to our cognitive constitution. This creates a phenomenal/noumenal distinction that involves causal transactions between empirical objects, which are mind-dependent phenomenalistic entities, and non-spatiotemporal noumenal grounds. This metaphysical picture violates the intelligible framework Strawson sees in the first face of Kant. The first face includes a

[22] Kant states, in the B-Preface of the first *Critique*, that the positive function of the critique of pure reason—i.e. the effect of its negative restriction of knowledge to experience—is to limit the pretensions of theoretical reason, so that it does not preclude the possibility of a practical extension of cognition to the supersensible objects of God, freedom, and immortality. *KdrV* Bxxix.
[23] Peter Strawson (1966): 4.
[24] *KdrV* B132.

principle of significance, whereby, Strawson thinks, there can be no meaningful[25] employment of the categories of the understanding apart from empirical conditions. In Strawson's judgment, Kant would have been better off pushing forward exclusively with his descriptive metaphysics, i.e. uncovering the fundamental principles immanent to experience that mark the parameters of what is intelligible to us. This descriptive metaphysics exposes the error of a Humean empiricism, which isolates individual impressions as epistemologically fundamental—i.e. it exposes the Myth of the Given. The Kantian, on the contrary, recognizes that these isolated impressions actually depend on a larger cognitive structure involving not only a "I think" unifying them, but also structures involving substances in mutual interaction in virtue of which isolated impressions can together form part of a single world and have an objective time-determination.[26]

Strawson legitimately identifies a tension in Kant. On the one hand, Kant's anti-Leibnizianism leads him to question and ultimately reject the rationalist project of establishing ultimate ontological features through the application of pure logical principles. We cannot endorse the attempt by Leibniz to develop a full demonstration of metaphysical reality by way the principle of contradiction, for instance. The Leibnizian moves immediately from the fact that something does not involve a contradiction, to the conclusion that it is possible. This privileging of the recognition by the intellect of a lack of contradiction makes the conditions of logic identical to the conditions of the

[25] Take note that Strawson says meaningful, and not justified or cognition-yielding. He is completely wrong to think that Kant is claiming that there is no meaningful use of the categories in abstraction from sensibility. Strawson here is conflating cognition with meaningfulness. Peter Strawson (1966): 4.
[26] See Kant's argument for the Third Analogy of Experience, in which he argues that objective simultaneity of events requires underlying substances in mutual interaction with one another. *KdrV* B257.

metaphysical structure of the world. But, Kant worries that this presumption of seamlessness from subjective rational principles to the metaphysical structure of the world leads to the validation of all sorts of empty fabrications by the architectonic mind—a mind inclined to system-building.[27]

What controls the overly fertile imagination of the rationalist metaphysician, limited only by the principle of contradiction, is sensible intuition. Sensible intuition adds another dimension to concepts in addition to the principle of contradiction: what is available to sensibility in the empirical domain further grounds metaphysical speculation in reality, and we can avoid empty fabrications.

We can articulate the tension identified by Strawson, between the descriptive metaphysics he finds valuable and the transcendent metaphysics he finds disastrous, in terms of a conflict between the spontaneity of the understanding, and that of reason. The understanding is spontaneous in that it brings to bears its own pure concepts on experience: the understanding is not defined in terms of passive affection by things, because it brings sensible representations under rules. Reason, according to Kant, has a higher form of spontaneity insofar as it does not merely provide a structure immanent to experience, but posits ideas that go beyond experience. Through its transcendent ideas, reason is able to distinguish the intelligible from the sensible words, and so to demarcate the limits of the domain within which the understanding might have cognition.[28]

A humanistic front in Kant scholarship generates a conflict between the spontaneity of the understanding and of reason, by absolutizing the spontaneity of the

[27] *Diss.* 2:416.
[28] *GMS* 4:452.

understanding. We see this absolutization poignantly in Strawson, in the restriction on meaningful employment of the categories to empirical content expressed by the principle of significance. In the humanistic interpretation, the absolutization of the understanding blocks the tendency of reason to extend itself beyond the sensible domain, because the understanding can only have significance in providing a conceptual structure that is immanent to experience.

The anti-Leibnizianism here, in its insistence on wedding cognition to sensible intuition, is the basis of the humanistic front in Kant scholarship that denies to Kant any commitment to a transcendent metaphysics.[29] Things in themselves, instead of being understood as an ontological ground underlying spatiotemporal appearances, have been recast as representing a merely formal and epistemological distinction between what conforms to conditions of intelligibility, and what does not.[30] Debunking things in themselves as metaphysical grounds takes out with it the idea of transcendental freedom as a real metaphysical cause, operating outside of time but nevertheless constraining the actions of the empirical character in time. Transcendental freedom is reframed as a mere standpoint rather than an actual metaphysical object. We recognize ourselves as free when we abstract from the epistemic conditions of experience, in which successive events in space-time are structured according to causal laws in accordance with the Second Analogy of Experience.[31] Humanistic interpretations of Kant do not need to make determinations about the non-ultimacy of spatiotemporal predicates. That is, identifying space and time as merely epistemic conditions apart from which we can have no

[29] Karl Ameriks (2000): 9.
[30] See, for instance, Henry Allison (2004).
[31] Henry Allison (2000): 221.

cognition does not make any ontological commitments about the relativity of space-time. Kant's contribution to the debate over the nature of space-time, in the humanistic interpretation, is an alternative to ontology rather than an ontological position in its own right.[32]

The humanistic front, then, rejects any metaphysics involving grounding relationships, i.e. one that characterizes reality in terms of fundamental substrata and their modes. A metaphysics involving the Kantian appearance-thing in itself distinction can distinguish between the thing as such, i.e. abstracted from any properties it takes on in relation to a human observer, and the relational properties things have when taken up by human sensibility. On this metaphysical view, we can conceive the thing in itself along the lines of a thin particular: that is, a fundamental substance abstracted from all relational properties, which is therefore the desired epistemic object of the pure ontologist, who identifies the intrinsic nature of things. This metaphysics would reject the idea that reality consists only in free-floating relational properties in space-time, ungrounded in a substance: slices of space-time that are always in relation to other slices in a contiguous location, conditioned events in causal relationships with their conditions, etc. The thin particular is the subject of this empirical content abstracted from all relation with human sensibility.[33] The language of thin particulars can be used to explain Kant's claim that appearances require something that appears.[34]

---

[32] Henry Allison (2004): 98.

[33] The terminology here including and related to thin particulars comes from David Armstrong (1999): 123-126. One can detect the idea of a thin/bare particular in the Paralogisms, in which Kant argues that knowledge of the subjecthood of the "I" underlying all thoughts cannot give us knowledge of the "I" in itself. *KrdV* A350.

[34] *KdrV* Bxxvii.

This metaphysics would be able to accommodate transcendental freedom as real form of causality, outside of empirical time. The underlying substance of the thin particular, abstracted from the relational properties in the empirical domain, could serve as the transcendental cause of the event-causal series in time. In this way, as I have noted, we could resolve some of the explanatory difficulties of agent-causation, by integrating the transcendental cause with the event-causal order, which can still function as reasons for action in the form of events.

The Disastrous Results of Setting the Spontaneity of the Understanding and that of Reason in Opposition to One Another

If the spontaneity of the understanding, however, is supposed to push out any metaphysics proposed by the spontaneity of practical reason, we end up with a metaphysics of agency that is forced to locate freedom within the continuous causal series of events in the empirical order. We are left, in other words, with a form of compatibilism that identifies freedom not with some real power to select among alternative possibilities, but a mere perspective. Strawson is famous also for developing an account of agency that is perspectival, and does not involve real causal powers.[35] He claims that, even if we establish scientifically that determinism is true, it would be psychologically impossible to refrain from engaging in reactive attitudes, in which we view one another as responsible persons instead of as mindless machines. We maintain our belief in personal responsibility, not because it is true, but because opting out of

[35] For Strawson, it is the perspective from which we entertain reactive attitudes, such as praise, blame, or resentment, that constitutes moral responsibility, independently of any theoretical justification. See the discussion in Gary Watson (1993): 120.

interpersonal relationships, and taking the objective attitude, in which we view other people as merely passive machines, would have too high of a psychological cost.[36]

However, this sort of metaphysics of agency has several serious vulnerabilities. First, the agency it posits, since it is subject to the deterministic causality of the empirical level cognized by the understanding, is causally irrelevant. The only causally controlling factors have to do with empirical events in the deterministic causal structure. Agency is merely phenomenological—the way we appear to ourselves—and this self-appearance is defeasible in light of an analysis of the causal structure which smothers our agency. There is a question, then, of whether maintaining reactive attitudes, in light of the absolutization of the spontaneity of the understanding which blocks out freedom as a real spontaneous power, is even appropriate. That is, if an individual is not truly the source of her action, as a sort of machine that merely processes external input according to built-in algorithms, the appropriateness of an interpersonal participatory attitude in relation to persons, rather than the objective managerial one we would take to the functioning of a machine, is questionable. We instinctively differentiate, to borrow an example from Strawson, between an individual who accidently steps on our foot and another who does it malevolently in our feeling of resentment peculiar to the latter. However, if both the accidental and the malevolent act are products of external input equally, a special participatory attitude in which we presuppose moral responsibility seems inappropriate to both.

[36] Peter Strawson (1962): 187-211.

Nevertheless, a standpoint account of rationality has emerged, in which we are able to take two contradictory positions, as long as we take them from different standpoints.[37] So, from the standpoint of our interpersonal relationships, we might view one another as really free, but from the standpoint of our scientific analysis, we might view one another as determined. Kant himself uses the language of standpoints, in his claim that the world of the pure understanding outside of appearances, is only a *standpoint* reason takes in thinking of itself as practical.[38]

However, Kant should not be misunderstood in this brief remark in the Groundwork, as claiming that our agency is a mere standpoint that has no objective realization as a genuine causal power. What he is saying in this passage is that practical reason does not intuit any objects in the intelligible domain. Objects of the will occur only in the empirical domain of sensibility. He writes, "If practical reason were to fetch in addition an *object of the will*, that is, a motive, from the world of understanding, then it would overstep its bounds…"[39] Only the formal condition of action exists in the world of the understanding, i.e. maxims that serve as laws of the will. We do not respond heteronomously in the world of the understanding, that is, in response to noumenal objects of inclination, but in response to normative judgments of reason. The lack of noumenal objects does not mean practical reason is a mere perspective and not a genuine causal power.

The standpoint theory, while un-Kantian, also has philosophical problems in its own right. It does not seem as though the presence of different standpoints absolves

[37] For instance, Christine Korsgaard (1989): pgs. 23-48.
[38] *GMS* 4:458.
[39] Ibid.

someone from irrationality in holding contradictory beliefs. Dana Nelkin points out that it is really impossible to believe something from a standpoint. If we only believe in our agency when taking a deliberative standpoint, but not from the scientific standpoint, then it seems as though we do not really believe it.[40] The standpoint would be more of a pragmatic heuristic for certain situations than something we actually believe about the world, since we cannot simply abandon our beliefs about the world according to considerations of convenience. If we truly believe in our freedom, then we believe in it *simpliciter*, i.e. not when it is useful, and this leads to a contradiction with a belief in determinism.

The conflict between the spontaneity of the understanding and that of reason, which leads to these flawed reductive accounts of agency from a standpoint, is really a manufactured one based on a misunderstanding of Kant. It is true that the understanding is restricted to sensible intuitions for cognition, but the limits of what is cognizable do not block the conditions of pure thought.[41] We can still think beyond possible experience, since the conditions of thinking are the mere absence of contradiction, and do not require sensible intuition.[42] The thought of the supersensible may not be justifiable from the theoretical perspective, but that does not make it meaningless.[43]

[40] She makes this point in Dana Nelkin (2000): 569.
[41] See *R* 3938, Cited in Guyer (1989): 59.
[42] See the difference between thinking and cognition in the B-Preface, at *KdrV* Bxxvi.
[43] Kant distinguishes between conditions of cognition and conditions of thinking, and claims that we are free to think with the categories independently of sensible representation, at *KdrV* B166.

Overview of Chapters

This dissertation, then, will aim to properly integrate the spontaneity of reason and that of the understanding, so that we can maintain a Kantian modesty in restricting the understanding to sensible intuition, while at the same time allowing practical reason to extend cognition into the domain of the supersensible. This integration will allow us to develop a metaphysics of agency in which absolute freedom can exist as a real causal power, instead of a causally irrelevant and irrational standpoint. In Chapter One of this dissertation, I go through various interpretations of transcendental idealism, with an eye to assessing what kind of metaphysical framework they can provide for the grounding relationship I think is essential to agent-causation. Of particular interest is the fact that epistemological interpretations of transcendental idealism seem to flounder with respect to grounding relationship between appearances and things in themselves. Making a merely epistemic distinction between what is in conformity with epistemic conditions and what is not has nothing to do with what is metaphysically conditioned and what is fundamental.[44]

In Chapter 2, I discuss Kant's evolution from a compatibilist in the same vein as Leibniz, to his rejection of compatibilism in the Critical period as the freedom of a mere turnspit.[45] I argue here that this fundamental shift in theory of freedom has to do with the

[44] Eric Watkins also argues that methodological interpretations struggle with the grounding thesis. His reasoning is that, since the methodological interpretations understand the transcendental distinction in terms of mutually exclusive perspectives, it is unable to formulate the grounding thesis. We cannot claim that things in themselves ground appearances, because this would require both that we both view things in conformity with, and out of conformity with, our epistemic conditions. Watkins (2005), 325.
[45] *KdpV* 5:97.

discovery of a new kind of grounding relationship, that both involves a diversification from the antecedently determining grounds of the *New Elucidation* and also an important modification of the Leibnizian notion of a *well-founded phenomenon*. The Leibnizian notion of an immanent ground of appearances makes spatiotemporal predicates completely transparent to things in themselves, since they are merely confused representations of the real internal structures of monads. It appears as though Kant thinks that the fact that temporal predicates are confused representations of monads that are reducible to the monadic structure, and not irreducible features of representations that need not correspond to their ground, means that temporality characterizes monads.[46] The fact that Leibnizian *spirituale* (i.e. monads) are subject to antecedent determination in time means for Kant that they are unable to be self-determining, since the past is necessarily out of our control.[47]

In Chapter 3, I will analyze the arguments of the first and third antinomy, and show how epistemological interpretations of TI are not able to properly establish a metaphysical foundationalism by which reason posits an unconditioned—though not in hypostasized form—which is supposed to bracket the conditioned cognitions of the understanding. Epistemological TI, in turning the appearance/thing in itself distinction into a formal one between what can and cannot be known, does not make the needed adjustments to avoid a schizophrenic tension between reason and the understanding. The schizophrenia between reason and the understanding, in which reason seeks a

[46] Some commentators claim that, since for each individual monad there is a complete concept of all its attributes, these attributes must be indexed to time in order to avoid contradiction, i.e. opposite characteristics occurring at the same time. Robert M. Adams (1994): 73.
[47] *KdpV* 5:97.

supersensible unconditioned but the understanding restricts cognition to what has empirical support, comes out most clearly in Kant's Antinomies. The antinomies arise because of the tendency of the understanding to absolutize its own cognitions. Since the antinomies presuppose a transcendental realism about space and time, in which these conditions of cognition are granted an absolute ontological status, reason, in the first antinomy, is forced to posit the unconditioned totality of the world in space and time. The contradictory results of this attempt lead to transcendental idealism, in which space and time are not absolute measures of reality but merely forms relative to human sensibility. The fact that space and time have a non-absolute status, such that they are *neither* finite or infinite features of an unconditioned totality, leaves room in the third antinomy for an intelligible space, outside of the causal series in time, for an unconditioned power of freedom.[48] However, if we remove the metaphysics of removing the unconditioned status from space-time, and instead attack the antinomies epistemologically, so that both the thesis and the antithesis represent a transcendental illusion into the nature of reality in itself, we fail to secure the possibility of this unconditioned power as a real object. We may defuse the antinomial conflict by refraining from all cosmological speculation as a "God's-eye" view[49], but this epistemic restriction does not guarantee that the causal structure of spatiotemporality does not block out the possibility of any genuine non-empirical cause.

In Chapter 4, I will turn my attention to three varieties of contemporary libertarian theories of free will[50]: event-causal, non-causal, and a form of agent-causal libertarianism

[48] *KdrV* A444/B472.
[49] The idea of a God's-eye perspective comes from Allison (2004): 27.
[50] Libertarians believe that free will is incompatible with determinism, and that it exists. Timpe (2013): 24.

that relies on a dubious notion of emergent properties.[51] I will show how limited and implausible these theories are, and account for these deficiencies in terms of a crude empiricism that attempts to ground empirical cognition on discrete impressions, while rejecting any structural metaphysics underlying these impressions involving substances and causal powers. Contemporary libertarianism attempts to ground its theory of free will entirely on discrete impressions in a neo-Humean way, so that it is predicated on observable events connected in a loose indeterministic fashion, without any metaphysical sub-structures. This approach leads to forms of agency in which the ultimate orientation of the will falls out from a chance process, and is not subject to agential control. The way out of the limited agency framed by an empiricist epistemology is a regressive Kantian methodology, in which empirical knowledge requires rationalist structures, in the form of substances and causal powers, in order to undergird the unity necessary to empirical knowledge. We see the importance of a priori structures particularly in the Analogies of Experience, in which Kant shows how time-determination of events (i.e. whether they are simultaneous or successive) can only have an objective status against the backdrop of substances and causal powers which provide the stability necessary for time-determination.[52] The very unity of time, in which we are able to posit different impressions as succeeding one another in the same time, depends on a substantial metaphysics that prevents our experience from fragmenting into new and independent time-frames.[53] Empirical cognition depends, then, not merely on certain cognitive

[51] Timothy O'Connor (2000): 114.
[52] *KdrV* A182/B225.
[53] *KdrV* A188/B231.

structures but also on the way experience is ontologically constituted[54], and this recognition can lead us out of the inadequate conception of agency based only on discrete empiricist impressions.

In Chapter 5, I develop in substantive detail an account of agent-causation using the metaphysics of transcendental idealism, and show how it is superior to the libertarian accounts I describe in Chapter 4, as well as other prominent agent-causal accounts. I also rebut charges against timeless agency, and also that we can rule out the possibility of agent-causation on the basis of empirical evidence that suggests continuous lawfulness.[55]

[54] Eric Watkins argues for the importance of ontological constitution to time-determination by tying the Analogies to Kant's pre-Critical metaphysics. See Watkins (2005): 215.

[55] I am referring to Derk Pereboom's objection that, since current empirical evidence indicates an uninterrupted lawfulness, the only way an agent-cause could exist would be if its free actions somehow coincide at every juncture with what happens according to empirical laws. Derk Pereboom (2001): 79. Pereboom makes a case against agent-causation again (though he does not accuse it of incoherence) in Derk Pereboom (2014): 52.

CHAPTER 1: TRANSCENDENTAL IDEALISM AND THE METAPHYSICS OF AGENCY

Our agency, while phenomenologically compelling, becomes problematic when subjected to philosophical reflection. Our sense of direct control over our actions clashes with our awareness of explanatory frameworks in which we can account for our behavior in terms of external influences—i.e. our upbringing, our genes, the company we keep, temperamental quirks that are inborn, etc.[56] The tension between an ability to be the source of one's own action, in an act of spontaneous causality, and a membership in the empirical order, in which all events are subject to antecedent events according to a causal law, is the subject of Kant's third antinomy.[57]

Transcendental idealism (hereafter abbreviated as TI) is essential in formulating Kant's metaphysics of agency.[58] In this chapter, I will assess the implications of different interpretations of TI for developing a metaphysics of agency. In section one, I will develop and assess a fundamental contrast in terms of a metaphysics of agency, which establishes a relationship between an act and an agent and also shows how the agent is to be integrated into the natural order, involving an event-causal metaphysics and an agent-causal approach involving substance causation. In section two, I will explain how TI is able to resolve a crisis in the event-causal model, in which a demand to account for the arising of events through previous events leads to an infinite regress in which no event is

[56] See the story of the hardened criminal Robert Alston Harris, seen from the perspective of both human freedom and external influences, in Derk Pereboom and Michael McKenna (2016): 3.
[57] We find the third antinomy at *KdrV* A446/B474.
[58] "For if appearances are things in themselves, then freedom cannot be saved." *KdrV* A336/B564.

ultimately explained, as well as a crisis involving the grounding of agential control. TI is able to resolve this crisis by introducing temporally indeterminate activities at the basis of the series of events at the empirical level. In section three, I will explain how the vision of TI's resolution of the crisis of event-causal models of agency is unacceptable and/or variously articulated in different interpretations of TI, and how there is a tension in interpretations of TI in emphasizing either empirical realism or metaphysical realism. In section four, I will discuss how different interpretations of TI adjudicate the tension between empirical and metaphysical realism by assessing these interpretations in terms of three core commitments of TI identified by Lucy Allais. In section five, I will provide close detail on specific representatives of epistemological interpretations of TI, and give a brief critique of each. In section six, I discuss metaphysical interpretations of TI, and provide an evaluation.

The Metaphysics of Agency: Event-Causal Vs. Agent-Causality

*Event-causal metaphysical frameworks***.** Several metaphysicians of agency in the late twentieth and early twenty first century have adopted an event-causal framework for agency, in virtue of which actions are events involving states of agents, rather than agents themselves.[59] Actions are distinguishable from non-action by way of a certain event-causal history, i.e. histories that involve states of the agent and not impersonal events detached from the agent. Donald Davidson, for instance, understands agency in terms of

[59] See A. Goldman (1970), and B. Enc (2003).

belief-desire combinations, structured in the form of a practical syllogism, in which the major premise establishes a desire, and the minor premise establishes the means the agent believes will achieve this desire.[60]

If agency can be established in virtue of a belief-desire combination, it is perfectly compatible with a larger naturalistic event-causal order beyond the agent. Action in an event-causal sequence is not rooted in some special unconditioned power unique in relation to other natural causes that merely follow prior causes. The introduction of substance-causation in virtue of which an agent determines itself independently of events, would represent an interruption of a naturalistic event-causal order. The idea of an event-causal order is a product of the heavy influence of Hume on contemporary philosophy, and it is significantly different from accounts of causation involving substances and their powers. Humean event-causal models break down reality into localized occurrences in space-time, and reject the idea of underlying substances that produce events in virtue of powers. For the Humean, a strict empiricism rejects the idea of a substance underlying the sensible qualities of an object—an unknown *something* in which colors, tastes, shape, etc. inhere.[61] Humean empiricism also breaks down the notion of power, claiming that causation consists merely in contiguity and succession of two events, and there is no impression corresponding to a faculty underlying events whereby a lawful connection can be established.[62] Laws of nature for the Humean are not based on underlying powers or natures that constrain the flow of events to certain structural regularities, but are instead

[60] Donald Davidson (1963): 691.
[61] David Hume (1949): Book I, Part I, Section VI.
[62] Ibid., Book I, Part II, Section XIV.

accidental generalizations based on past accumulation of events themselves.[63] Events are indivisible points occurring at discrete instants, and cannot be assimilated into enduring metaphysical structures involving substances and powers. Even though the Humean denies any causal power uniting events, the Humean can still develop an event-causal metaphysics involving constant co-occurrence of types of psychological motivations and types of action.[64]

*Agent-causal metaphysical frameworks***.** In opposition to Humean event-causal metaphysics, there are metaphysical frameworks for agency that involve substance-causation. Substance-causation is non-occurrent, insofar as it does not consist in an event. Substance-causation, therefore, can stop a regress in event-causal frameworks. Events arise at a particular time—e.g. the ball *flew* through the air, *struck* the window, and the window *broke*. Every arising of an event requires some prior event to account for its arising.[65] The way to stop this infinite regress of events is to introduce substance-causation, by which an event is caused in virtue of an activity of a substance, instead of a temporally localized event. Kant in the argument for the Second Analogy claims that, in order for an event to be temporally localized, it must follow from a preceding event according to a causal law, which makes for an objective temporal sequence.[66] At the same time, though, Kant demands that there is an "ultimate subject of the changeable,"

[63] For the contrast between the Humean and the Kantian account of laws of nature, see Eric Watkins (2005): 345-6.
[64] Hume (1949): 2.3.1.14
[65] Kant claims, in the antithesis argument of the third antinomy, that the unity of experience requires that every event be connected to a preceding event according to a causal law. *KdrV* A444/B472.
[66] "The principle of sufficient reason is the ground of possible experience, namely the objective cognition of appearances with regard to their relation in the successive series of time." *KdrV* A201/B246.

that is, an underlying ground of successive events which itself does not change. Kant claims that *actions* unite this ultimate subject of the changeable to events which change.[67]

There is something intuitively compelling in the account of an action as a primary ground of change. If we assimilate action merely to the changeable, i.e. the series of occurrent events following one another according to causal laws, then action becomes, paradoxically, something passive. If action just follows from a preceding event, then what differentiates it from mere reflexes, like blinking? Alexander Baumgarten, in whose metaphysics Kant was thoroughly immersed, defined action in terms of control over the sufficient cause of an event.[68] Being in control of the sufficient cause of an event means that action is not caused by some preceding event external to the agent. If all we have are events in our toolkit for developing a metaphysics of agency, there is no way to root an action in an agent. We might appeal to an event by which the agent causes an action—some exertion on the part of the agent to produce an action—but, then, either this event arises of itself, or there has to be another event by which the agent controls it. With events, one faces either an arbitrary arising, an infinite regress of events by which an agent causes herself to cause an action[69], or a connection to alien events external to the agent.

---

[67] *KdrV* A205/B250. Also, in *ND*, Kant does not appeal to changing grounds to appeal changing determinations of substances, since changing grounds would simply push back the problem of explaining change. Watkins (2005): 116. This demand for an unchanging subject is the basis for a criticism of Leibniz's attempt to account for change solely in terms of derivative, and not primitive, forces. Watkins (2005): 127.

[68] Alexander Baumgarten (2013): No. 704. Knutzen, Kant's teacher, has a similar definition, cited in Watkins (2005): 57.

[69] It is Roderick Chisholm who claims that an agent-causal theory must account for how the event of agent-causation is controlled by the agent. If we appeal to another event to account for how the agent controls the event of agent-causation, then we are led into an infinite nesting of agent-causal events. Roderick Chisholm (1971): 40.

In the argument for the Second Analogy of Experience, Kant points to the necessity, for temporally determinate changes in time, of a temporally indeterminate ground—he steps out of the event-causal framework, by appealing to temporal indeterminacy, and so avoids the three alternatives I just identified. Kant accepts that the objective arising of events in time requires a causal law that situates an event as truly successive to a previous event.[70] A temporally indeterminate ground[71], in which no event arises, is therefore necessary to stop an infinite regress of attempts to account for the arising of events. The crisis of the event-causal model is evident in the Third Antinomy, in which an ambiguity in the principle of sufficient reason leads to a contradiction. On the one hand, we might interpret PSR as requiring *completeness* in one's explanation, and so a series of conditions must bottom out in something unconditioned, i.e. an uncaused subject of action. On the other hand, we might interpret PSR as requiring that every event must follow from a previous event according to a causal law—this is a universality requirement stemming from PSR. Of course, an unconditioned cause would not follow from a causal law, and so the universality and completeness requirements of PSR are in conflict.[72]

However, we can reconcile this tension in PSR with a metaphysics of temporally indeterminate activity underlying events: substances able to act, through a determination of their own states, do not need an external event to account for the activity in their own power. The activity of the substance does not arise in time, but is ongoing, and so there is

[70] This is the argument of the second Analogy of Experience. *KdrV* B233.
[71] The idea of a temporally indeterminate ground comes from *ND,* where Kant claims that a substance is immutable in itself. Changing determinations can only come through the introduction of new grounds from interaction with other substances. *ND* 1:411.
[72] Henry Allison eloquently articulates this contradiction as involving a completeness requirement and a universality requirement in PSR. Henry Allison (1990): 22.

no need to account for it in terms of another event that occurs at a specific time. The dual requirements of PSR, for both completeness and universality, can be satisfied only if we break free from the event-causal model. The activity of temporally indeterminate substances underlying events can satisfy the completeness requirement, without violating the universality requirement, since what is temporally indeterminate does not need to be accounted for in terms of a preceding event.

Transcendental Idealism Resolves the Crisis of the Event-Causal Model

The crisis of the event-causal model is twofold. First, it demands an accounting for each event in terms of a previous event, but this leads to an infinite regress in which no event can be satisfactorily explained. Second, event-causal models do not allow us to root action in an agent, since we are faced with a trilemma of events that arise by chance, deterministic events, or an infinite regress of attempts by the agent to control its own states through more events. We can introduce some temporally indeterminate activity to resolve this crisis, but an ontology that consists exclusively in events cannot accommodate such an activity. Transcendental idealism can allow us, by relativizing the event-causal order in the empirical time-series as appearances dependent on the human subject, to introduce the underlying substances that ground the time-series and can complete the explanatory series.

Action, for Kant, incorporates a notion of unconditioned causality that cannot be completely empirical. The appeal to a non-empirical element comes from the fact that the

understanding, the concepts of which necessarily underly all intuitions as an a priori structure[73], requires connectivity according to the causal law for objective time-determination.[74] This kind of connectivity is incompatible with an unconditioned ground, and so transcendental freedom, if it is to exist at all, must exist on an intelligible level, i.e. a level that we can only think of but not have cognition of in experience.[75]

The causal law in appearances, which establishes an event-causal framework in which each event follows from a preceding one in objective succession, cannot characterize all causality, if an unconditioned cause is to be possible. Transcendental idealism can ensure the lack of sufficiency by appearances, by establishing that appearances are in themselves indeterminate[76], and require, in addition to the causal law of successive events, an underlying ground.

This sort of metaphysical structure is a kind of metaphysical foundationalism, in which the event-causal structure of appearances, in which successive events follow from an antecedently determining ground in time[77], itself is conditioned by an underlying ground, in which an unconditioned causality may exist. Metaphysical foundationalism is distinct from metaphysical infinitism, in which there is no unconditioned, only an endless conditioned-condition series, and metaphysical coherentism, in which conditioned entities interact in reciprocal feedback loops. Metaphysical foundationalism, on the contrary, presents well-founded levels of reality, that are based on what is unconditioned

[73] This is the conclusion of the Transcendental Deduction. *KdrV* B166.
[74] This kind of connectivity embodies the law of inertia, whereby only extrinsic causes can move an object out of its current state. Discussed in Watkins (2005): 69.
[75] Thinking is severed from sensible intuitions, whereas cognition requires it. *KdrV* B166n.
[76] Kant says that appearances are the undetermined objects of empirical intuition. *KdrV* A20/B34.
[77] Kant defines an antecedently determining ground in his *ND*: "An *antecedently* determining ground is one, the concept of which precedes that which is determined." Immanuel Kant (2002): 1:392.

and ontologically independent. The metaphysical foundationalism I am envisioning is not founded on spatiotemporal particulars—i.e. individual events in the Humean universe of loosely associated occurrences lacking underlying substances. Instead, the metaphysical foundationalism gives priority to a purely intelligible level which represents an objective architecture of being that is independent of the spatiotemporal perspective peculiar to human sensibility. The intelligible structure is the intellectual synthesis of the B-Deduction, in which the categories are used purely to think of objects in general, and so in abstraction from the conditions of human sensibility.[78] The intelligible structure is not controlled by antecedent events in time, since it is independent of the temporality of human sensibility, and so it can satisfy the completeness requirement of PSR.[79] An agent is also to able to establish control of its action by serving as its intelligible ground, and it does not therefore have to establish such control by way of an event, which itself would require yet another event to establish control, leading to an infinite regress. The intelligible ground, since it avoids the infinite regress of events, establishes *ultimate responsibility* for action, i.e. the source of its sufficient reason.[80]

[78] Kant defines the intellectual synthesis (*synthesis intellectualis*) at *KdrV* B150.
[79] The difference in temporality between the intelligible and empirical levels highlights an important difference between Kant and the Leibnizians, who thought that things in themselves were parts of appearances, such that appearances were merely a confused and indistinct version of things in themselves. For the Kantian, things in themselves are a distinct substrate of appearances, not parts of them. See Rae Langton on the Kant-Eberhard controversy, in Rae Langton (1998): 192.
[80] The phrase comes from Robert Kane, cited in Timpe (2013): 58.

Different Interpretations of TI and the Vision of the Resolution of the Event-Causal Crisis

The glaring problem, though, is that the sort of metaphysical foundationalism that addresses the crisis of event-causal agency is unacceptable to many interpreters of Kant. There are major readings of Kant in which any talk of an intelligible metaphysical structure bracketing empirical reality is either completely irrelevant, or even incoherent. Even those who seek to incorporate an intelligible metaphysical structure differ amongst themselves. A major test for an interpretation of idealism is to what extent it can incorporate the two forms of realism we see in Kant. That is, we need to inquire into the implications an interpretation of TI has for an *empirical* realism about spatiotemporal objects in experience, and a *metaphysical* realism about objects of which we never could have cognition, at least through theoretical reason, because they transcend the boundaries of possible experience.[81] While empirical realism involves an application of the pure concepts of the understanding to a sensible object in space and time, metaphysical realism involves the thought of an object *in general*, in which pure categories are applied independently of the limitations of sensible intuition.

The integration of empirical and metaphysical realism takes the form of two kinds of legislation, which correspond to the resolution of the tension of PSR I just discussed. On the one hand, there is legislation of the understanding, in virtue of which it imparts pure concepts that provide a structure to nature. This structure is event-causal, such that

[81] Peter Strawson discusses these two forms of realism with great clarity in Peter Strawson (1994): 168-70.

causal laws structure relationships between events, so that the lawful unity of nature does not permit a lawless substance-cause able to determine itself independently of any contextual events.[82] On the other hand, there is the legislation of reason, which posits a concept of freedom that has absolute spontaneity and so cannot exist in the causally structured empirical time-series. The way to reconcile these two apparently contradictory forms of legislation is through the idea of the supersensible: the empirical series is qualified as a mere appearance, relative to human sensibility, opening up a domain for an intelligible order not subject to the conditions of human experience.[83] However, this reconciliation is not without controversy. There are those who think that the presence of a more metaphysically fundamental domain underlying appearances makes these appearances merely subjective phenomena. If appearances are not metaphysically fundamental, according to some commentators, then they are phenomenalistic entities, i.e. private psychological entities that provide access only to the contents of the mind but not to reality.[84] If, on the other hand, we disregard Kant's metaphysical realism, expressed in the legislation of reason in the concept of unconditioned freedom, then Kant's philosophy can never address the crisis of the Enlightenment. This crisis lies in the fact that reason unfettered by authority led to the mechanistic conception of physics,

[82] See the Remark on the Antithesis of the third antinomy at *KdrV* A451/B479.
[83] Kant discusses these two forms of legislation and their reconciliation through the supersensible in *KdU* 5:175.
[84] This is a very uncharitable commentary on the metaphysical interpretation of TI, since the fact that something is not metaphysically absolute does not means that it loses all objectivity and becomes a private phenomenalistic entity reducible to psychological states. Karl Ameriks explains this in Karl Ameriks (2012): 81-2.

which seems to preclude free will. The theoretical autonomy exercised by reason in science ironically reveals the absence of practical autonomy.[85]

Three Axes of TI and A Spectrum of Different Views on Them

In her recent and important book, Lucy Allais identifies three central components of TI[86]:

1) A distinction between appearances (*Erscheinungen*) and things in themselves (*Dinge an sich*).[87]
2) The idea that the spatio-temporal aspect of objects is a *mere* appearance or *mere* representation that is mind-dependent in some way. Kant makes clear that spatiality is a necessary condition of the appearance of objects, but not of their existence.[88] Similarly, Kant argues that, if we think of objects independently of our sensibility (i.e. in themselves), "then time is nothing."[89]

[85] See Kant's Preface in the B-edition, where he discusses how a critical distinction between sensible and intellectual domains allows us to integrate belief in supersensible objects with the mechanism of nature. *KdrV* Bxxviii.
[86] Lucy Allais (2015): 4.
[87] "Yet the reservation must also be well noted, that even if we cannot cognize these same objects as things in themselves, we at least must be able to think them as things in themselves. For otherwise there would follow the absurd proposition that there is an appearance without anything that appears." *KdrV* Bxxvi. Here it seems as though things in themselves represent a ground of things that appear. We cannot cognize this ground, but we have to think it because appearances by their nature require a ground.
[88] "The proposition, 'All things are next to one another in space' is valid only under the limitation that these things be taken as objects of our sensible intuition." *KdrV* A27/B43.
[89] "We cannot say that all things are in time, because with the concept of things in general abstraction is made from every kind of intuition of them, but that is the real condition under which time belongs to the representation of objects." *KdrV* A35/B51.

3) The claim that we cannot have cognition (*Erkenntnis*) of things as they are in themselves. "What may be the case with objects in themselves and abstracted from all this receptivity of our sensibility remains entirely unknown to us."[90]

Of these three axes of TI, I think the least controversial is axis three. No commentator thinks that Kant holds we can have cognition of things in themselves through theoretical reason. Attributing to Kant the ability to have metaphysical knowledge independently of experience, through purely rational means, would be to miss a fundamental distinction Kant is making between himself and the Leibnizian-Wolffian tradition. Leibnizian rationalists like Johann Eberhard attacked Kant for his doctrine of *transcendental illusion*. Transcendental illusion consists in the tendency to take subjective principles of one's reason as capable of establishing metaphysical principles of objective reality.[91] Leibnizian rationalists claim that mere deduction from principles of reason, like the principle of sufficient reason and the principle of contradiction, could lead even to the completion of metaphysics.[92] For Kant, cognition requires concepts *and* intuitions.[93]

That said, there is a controversy over whether Kant attributes to us the ability to have cognition of supersensible objects *through practical reason*. For instance, Lara Ostaric adjudicates between realist and anti-realist interpretations of Kant's notion of moral *Glaube*, which is a rational assent to matters related to the moral life.[94] Realist interpretations of the objects of moral *Glaube*—i.e. God, the immortality of the soul, and absolute freedom—conclude that belief in these objects corresponds to a real and given

[90] *KdrV* A42/B59.
[91] *KdrV* A297/B353.
[92] Henry Allison (1973): 17.
[93] "Thoughts without content are empty, intuitions without concepts are blind." *KdrV* A51/B75.
[94] Lara Ostaric (1999): 91-118.

object. Anti-realist interpretations of the objects of moral *Glaube*, on the other hand, view them as mere presuppositions of practical activity that do not correspond to real and given objects. In sum, there is a debate about whether Kant understands experience-transcendent objects as really existing apart from the mind, as a metaphysical realist, or whether experience-transcendent objects, like freedom, are merely part of a conceptual framework of a given perspective that do not represent some metaphysical fact.[95] Opposition to metaphysical realism emerges from an opposition to a verification-transcendent conception of truth, that allows for truth even beyond the limits of our ability to verify it.[96]

With respect to axis one, which distinguishes between appearances and things in themselves, there is a vigorous debate over whether this distinction is epistemological or metaphysical. The question here is whether or not the distinction between things in themselves and appearances corresponds to one between two ontologically distinct things or aspects of the same thing, or between two different perspectives on the same thing. Sometimes Kant sounds as though he is committed to the idea that appearances are mental states, and things in themselves are extra-mental objects, and so appearances and things in themselves are ontologically distinct.[97] Commentators like James Van Cleve

[95] There is a long tradition of anti-realist critiques of metaphysical realism. See Rudolph Carnap for an emphasis on relativity to a linguistic context. Rudolf Carnap, "Replies and Systematic Expositions," *The Philosophy of Rudolf Carnap*, ed. by Paul Schlipp, (La Salle: Open Court, 1963): 871. Henry Allison is a Kant commentator who refers to Michael Dummett and claims that freedom exists only from the practical perspective, and the idea that it might exist as a metaphysical fact is a transcendental illusion. Henry Allison (2004): 49.

[96] One can find an example of this neo-verificationist opposition to any experience-transcendent truth in Michael Dummett (1978).

[97] *KdrV* A490-1/B518-9.

attribute to a Kant a heavy *ontological* idealism about spatiotemporal objects in the empirical domain, which reduces them to mental objects.[98]

Other commentators attribute to Kant an *epistemological* idealism, which attributes a mind-independent existence to spatiotemporal objects but recognizes that our cognition of objects is informed by our own subjective cognitive structures.[99] Commentators who focus on epistemological idealism may be concerned with showing Kant's rejection of the Myth of the Given, that is, the idea that what is immediately given to consciousness—e.g. sensations, or impressions—can by itself constitute knowledge. Kant, of course, unpacks a whole process of synthesis in virtue of which a priori cognitive structures take up raw sensations in order for cognition to occur.[100] The exposure of an a priori component to knowledge that shows the error of classical empiricism in its focus only on isolated and fragmentary impressions is what Strawson, for instance, finds to be of lasting value in the First Critique.[101]

Reading Kant as exclusively an epistemological idealist leads to a transcendental distinction between appearances and things in themselves that is merely formal, rather than ontological. The transcendental distinction in epistemological versions of TI has to do with a distinction between what conforms to epistemic conditions and is cognizable, and what does not conform and so is uncognizable. The idea that things in themselves are epistemically opaque does not imply any metaphysical realism about the existence of

[98] James Van Cleve (1999).
[99] Josiah Royce makes a distinction between ontological and epistemological idealism in Josiah Royce (1892): xii-xiv.
[100] The Myth of the Given is an empiricist foundationalism. Eric Watkins (2012): 313.
[101] Peter Strawson (1966): 8.

supersensible objects.[102] The transcendental distinction in epistemological TI does not consist in two objects, or two aspects of objects, but in two perspectives on the same object. Epistemological TI therefore adopts the literal meaning—if not the spirit—of Kant's statement that the *Critique* "teaches us to name the object in a twofold significance, namely as appearance or as thing in itself."[103] While this comment is compatible with a merely formal appearance/thing in itself distinction, it is not *in*compatible with an ontological appearance/thing in itself distinction, since an object could have a twofold significance in terms of its metaphysical aspects as well as in terms of a conceptual distinction.

Axis two, which has to do with the mind-dependence of spatiotemporality, yields a range of interpretations that correspond to what extent one attributes to Kant an epistemological or an ontological idealism. Those who attribute to Kant a merely epistemological idealism will emphasize the epistemic role of space and time, claiming only that we can have cognition within space-time, but not that space-time is inferior in terms of its fundamentality with respect to an intelligible non-spatiotemporal level.[104] On the other hand, there are passages in the Transcendental Aesthetic in which Kant is very clear in defending a certain metaphysical idealism about space-time. He notes that, if we were to remove our subjectivity, space and time themselves would disappear.[105]

---

[102] Peter Strawson proposes a weak version of TI that merely establishes a principle of significance, to the effect that we cannot meaningfully apply concepts apart from empirical criteria. Peter Strawson (1966): 174.

[103] *KdrV* Bxxvii.

[104] Henry Allison, for instance, thinks that, in the disputes about the nature of space-time, Kant decides to avoid ontological questions about space-time altogether. Instead, he emphasizes the epistemic role of space-time Henry Allison (2004): 98.

[105] *KdrV* A42/B59.

The metaphysics of space-time are important with respect to the possibility of noumenal objects. The mere designation of space-time as an epistemic condition does not preclude the possibility that all of reality is in space-time. If all of reality is in space-time, then there can be no noumenal objects grasped by an intellect apart from the conditions of sensible cognition. Of course, the metaphysics of space-time are very impactful for our theory of agency. Kant notes that, if time is an absolute metaphysical structure, human beings would be nothing more than marionettes.[106] One of the possible threats to human freedom is the fact that we are created by God and so ontologically dependent.[107] The scenario in which human freedom would be destroyed is one in which temporality characterizes both God and humans with respect to their very existence, and not merely their appearance. In this case, human actions would be the mere effects of God's activity, and we would have Spinozism. On the contrary, if temporality is merely a *way of representing* human action and not a condition of being itself, we can delegate God's act of creation to the domain of things in themselves. God creates the ontology of human beings, not the way they appear. A creation of mere appearance would not really be a creation, since it would not affect the underlying ontology of the thing. God's act of creation therefore does not entail a subjection to the causal law in time, rooted in His will, because this subjection is a mere appearance that has nothing to do with God's act of creation.[108]

[106] *KdpV* 5:101.
[107] Kant considers different threats to freedom throughout his career. In his metaphysical lectures, the main threat is our dependence on God. In the first *Critique*, it is natural causality. Henry Allison (1990): 62.
[108] *KdpV* 5:102.

In his *Religion*, Kant talks about a propensity to evil that acts as a sort of super-maxim, that provides a formal structure for every act. As an umbrella maxim that covers every action in time, the propensity to evil has to originate outside of time.[109] The way we interpret such doctrines depends greatly on our original understanding of TI. If we opt for a reductive, epistemological understanding of the relationship between appearances and things in themselves, they are only distinct insofar as they are different *ways of considering* things. Within this perspective, freedom from ontological determination by the will of God, or the ability to choose a highest maxim that underlies every action, would represent a mere mode of self-consideration, and not a real supersensible power. We might use epistemological TI, then, to formulate a dual standpoint account of freedom, along the lines of Christine Korsgaard. Temporality, and its related determinism, may characterize us from a theoretical third-person perspective, but we can nevertheless claim a kind of autonomy from a deliberative standpoint, in which we abstract from the determinism of temporality, and act in light of reasons. Freedom is not a metaphysical property of the self, but a mere mode of consideration we take up from a deliberative standpoint.[110]

It is worth noting that the dual standpoint theory of freedom is a complete reversal of the kind of metaphysical foundationalism I sketched earlier. In this version of metaphysical foundationalism, practical reason provides the keystone concept of the architectonic of reason, in the form of an unconditioned causality—i.e. absolute freedom—that theoretical reason could only think about, but not cognize.[111] Here,

[109] *R* 6:32.
[110] Christine Korsgaard (1996): 94-8.
[111] *KdpV* 5:4.

practical reason provides a necessary complement to the temporally determined cognitions of theoretical reason, in virtue of which the empirical character of the self in time has an underlying ground that provides a form for the empirical character.[112] In Korsgaard's dual standpoint theory, it seems as though the level of empirical causation is foundational, and the practical standpoint inserts itself not as a metaphysical complement, but a mere anti-realist hypothesis underlying one's practical deliberation.

But, Kant takes a very different strategy in regards to our deliberative standpoint. Practical reason has primacy for Kant, which means that what is presupposed through our practical agency is able to add an increment to theoretical reason, which by itself is not able to cognize the supersensible. What is theoretically undecidable—e.g. whether or not we are free—can be decided in virtue of practical reason, which supersedes the epistemic limitations of theoretical reason.[113] In Korsgaard, since there is no apparent escape from the deterministic metaphysics established by theoretical reason—i.e. it has not been "disciplined"[114]—the rational agency posited by practical reason comes across as a kind of quirk. Reason in Korsgaard becomes schizophrenic, positing incompatible visions of itself, or, at best, compartmentalized, so that it is unable to fully unify itself. As Dana Nelkin has pointed out, without some ontological space whereby practical reason can posit freedom as a metaphysical reality, rather than a mere phenomenology of practical deliberation, belief that one is both free and determined is irrational.[115]

---

[112] I will develop the relationship between the intelligible character and empirical character more in coming chapters. For now, I will note that the role of the intelligible character has a role similar to substantial forms in Leibniz, which provide an intelligible structure to otherwise indeterminate matter. See Robert M. Adams (1994): 317.

[113] *KdpV* 5:121.

[114] See the Discipline of Pure Reason, *KdrV* A708/B736.

[115] Dana Nelkin (2000): 567.

Epistemological TI

Epistemological interpretations arise in response to a demand to reframe TI in such a way as to avoid accusations of incoherence that dog metaphysical versions of TI. For instance, metaphysical versions of TI subscribe to what is known as *noumenal affection*. Noumenal affection is the idea that the experience of spatiotemporal objects is generated via the affection of the mind by things in themselves that are unknowable.[116] The obvious objection to this phenomenon is that we cannot both claim that things in themselves cause experience, and that we cannot know things in themselves.[117]

Having attributed noumenal affection to Kant, Peter Strawson then goes on to judge this doctrine to be inconsistent with another doctrine he finds central to TI, the principle of significance. This principle is the rule that we cannot use concepts in propositions that represent knowledge claims unless these concepts have some empirical application. The principle of significance rules out the affirmation of noumenal affection, which establishes a causal relationship between that which is empirical and that which is non-empirical.[118]

Some commentators who interpret TI as involving the metaphysics of noumenal affection go on to reject it, and extract from the Critical philosophy what is still salvageable.[119] The approach of the epistemological interpretation is to reframe TI so that

[116] Desmond Hogan (2009): 502.
[117] We find this objection in Jacobi (1815): vol. 2, 303.
[118] Peter Strawson (1966): 235-245.
[119] Henry Allison refers to this approach to Kant scholarship as the separability thesis. Henry Allison (2004): 6. Guyer endorses the idea that TI can be dismissed, while the transcendental theory of experience in the Analogies and the Refutation of Idealism can be retained, in Guyer (1987): 335.

it no longer has the incoherent metaphysics of noumenal affection, and so TI can be retained as a valuable philosophical contribution. Instead of tying the appearance-thing in itself distinction to two objects—one empirical and spatiotemporal, the other noumenal and non-spatiotemporal—epistemological interpretations tie the transcendental distinction to two ways of considering objects.[120] Not all epistemological interpretations grant even a modest negative, or epistemic role, for things in themselves—e.g. a certain disciplinary role in checking the pretensions of reason to go beyond sensible conditions. Some epistemological interpretation deny the very intelligibility of the thought of a thing independent of the conditions for knowing it.[121]

Arthur Melnick, for instance, thinks that the concept of an object is not an ontological notion, having to do with some structure in reality that is independent of any perceiver. Instead, the concept of an object is an epistemic notion by which a perceiving subject organizes its experience. The Transcendental Deduction is centrally concerned with establishing that the pure concepts of the understanding apply to all sensible intuitions as their a priori structure—i.e. it rejects the idea that there is a misfit between our pure concepts and the world of sensible objects.[122] Objecthood, which is constituted according to the pure concepts of the understanding, is for Melnick is an essentially epistemic notion having to do with the way an individual self-consciousness synthesizes its own experience. Just as an object is an object only in relation to cognitive structures, a pawn is only a pawn in relation to the rules of chess.[123] Melnick's version of Kant is an

[120] See, for instance, Graham Bird (1962): 36ff.
[121] Karl Ameriks (2003): 69.
[122] *KdrV* A85/B117. The specter TD tries to exorcize is that of cognitive emptiness, i.e. the possibility that the most fundamental categories of our understanding distort our perception of the world rather than objectively apply to it. Watkins (2005): 188.
[123] Melnick brings up this analogy in a review of Allison's *KTI*, in Arthur Melnick (1995): 134-136.

epistemological idealist, insofar as he identifies idealism with a priori cognitive structures we need to cognize an object, but not an ontological idealist for whom the existence of objects is somehow dependent on the mind.

A thing in itself, which is transcendentally outside of the subject and its entire cognitive apparatus, is concept of an object which is completely severed from the very conceptual structure in virtue of which we understand objecthood. The idea of a thing in itself is therefore an absurdity.[124]

Critique of Melnick. Melnick's route to epistemic humility about things in themselves is short.[125] Objecthood is essentially related to cognitive structures articulated in the Transcendental Deduction, and so any notion of objecthood that is independent of these structures is incoherent and therefore unknowable. Short arguments to epistemic humility reach humility through a simple reflection on non-specific features of our cognition, such as representation as such, or receptivity as such[126], and then conclude that, since we have representations, or our intellects are passively receptive to impressions, we cannot know reality in itself. However, it appears as though Kant took a longer route to humility: this route involves a metaphysics of space and time in the

[124] Arthur Melnick (1973): 133-154. For a commentator with a view similar to Melnick's, see Jonathan Bennett (1974): 52. Bennett attributes to Kant a concept empiricism, whereby only assertions about possible experience are intelligible for us. The notion of a thing in itself outside of all possible experience is not even a coherent notion.

[125] I am referring here to Karl Ameriks's distinction between short and long arguments to humility, in Karl Ameriks (2003): 136-7.

[126] For instance, Rae Langton distinguishes between appearances and things in themselves in terms of relational qualities, which incorporate the receptivity of our minds in relationship to objects, and intrinsic qualities, which are independent of any relationship with our receptive intellects. Langton's distinction allows her to present Kant as a scientific realist, since appearances can stand in a relationship to receptivity even if they are microscopic or extremely distant from us. The appearance/thing in itself distinction in Langton has nothing to do with an idealism about space and time. Rae Langton (1998): 189.

Transcendental Aesthetic[127], and a defense of the non-totality of the world in the First Antinomy. These arguments establish that space and time are conditioned realities specific to the sensibility of the human subject, and so the world in space and time cannot exist as a totality independently of the human subject. The Copernican Turn in the Transcendental Deduction, whereby objecthood is tied to epistemic capacities of the human subject, has to be complemented by arguments for the ideality of space and time and their specific characterization of human sensibility.[128] The union between pure concepts of the understanding and intuition can only take place in space and time, since we can only have intuitions that are mediated through the pure forms of sensibility.[129]

When we adopt the long argument to humility, it becomes clear how the thought of things in themselves would not be incoherent. A thing in itself would be a thing independent of the ideality of the spatiotemporality of our kind of intuition. A non-spatiotemporal object may not be a possible object of experience, but it is not incoherent, such as the representation of an object that is independent of all representation. If we adopt the short argument to humility, and claim we cannot know things in themselves because they are independent of representation as such, we find that the very attempt of formulating this representation-independent thing is a form of representation, and so the whole notion degenerates into paradox. This sort of critique of metaphysical realism, which manifests itself in Melnick, conflates objecthood with the structures of

[127] The metaphysics of space and time is also very important in Kant's pre-*Critical Inaugural Dissertion. Diss* 2:403.
[128] Ironically, Melnick himself, in a review of Henry Allison's *KTI*, notes that the Copernican turn in epistemology, whereby we know things only in virtue of certain epistemic conditions, can be accepted also by transcendental realists who do not accept the non-totality of the world or Kant's metaphysics of space and time. Kant is giving us more than a mere description of an epistemological idealism in virtue of which we know things under certain conditions. Arthur Melnick (1985): 135.
[129] *KdrV* A89/B122.

representation, and fails to take into account the idea that spatiotemporal objecthood is a contingent form of objecthood subject to the form of human sensibility. We can think of a thing in itself beyond this form of our sensibility, through the pure use of the categories. In this pure application of the categories, we do not abstract from representation as such, but from a contingent kind of representation, i.e. spatiotemporal representation.

Clearly, views like those of Melnick's fare well with respect to a realism about spatiotemporal objects. The claim that we know objects in virtue of certain epistemic structures does not threaten the mind-independence of such objects. Kant is an empirical realist in the sense that he recognizes an empirical thing in itself, which marks an appearance/reality distinction within appearances. Kant distinguishes between a rose as an empirical thing in itself, and the way it appears to different people subjectively constituted in different ways.[130] Kant can also be understood as a scientific realist in light of the fact that he attributes objective reality to unobservable entities posited in scientific theories. For instance, Kant thinks we can cognize the presence of a magnetic force penetrating all bodies from the perception of attracted iron filings, even though our unaided senses are too crude to directly perceive this force.[131]

Such a commitment to empirical realism seems incompatible with an ontological idealism. If appearances are mind-dependent, then it seems as though it would be impossible for them to exist apart from observation, and for us to be able to distinguish

[130] *KdrV* A30/B45.

[131] *KdrV* A226/B274. Rae Langton recognizes that Kant includes what is unobservable in possible experience. Unobservable grounds are still possible experiences for Kant insofar as we can stand in receptive relationships with them. Things in themselves, for Kant, are still more fundamental, insofar as they are not even objects of possible experience. Leibnizians like Eberhard conflated unobservable grounds with things in themselves. Rae Langton (1998): 197.

between objects and subjective features of things. However, there are ways of framing an ontological idealism that allows us to continue to preserve the empirical thing in itself associated with empirical realism. We do not need to completely banish the intelligibility of the thing in itself to preserve empirical realism. Ontological idealism does not have to be a kind of identity phenomenalism, in which empirical objects are reduced to mental states.[132] Kantian idealism ascribes the spatiotemporal form of objects of objects to our way of knowing them, but nevertheless there is no denial of mind-independent objects.[133] We could claim that appearances are ontologically ideal in the sense that the spatiotemporal form of objects characterizes a specific kind of sensibility, and is not an absolute/unconditional measure of reality.[134] The *objects* in space and time are not ideal. So, the empirical object is not reduced to a private mental entity, but it there is an element of genuine ontological idealism insofar as they spatiotemporal form of objects is metaphysically non-absolute. "…In so far as [phenomena] are sensory concepts or apprehensions, they are, as things caused, witnesses to the *presence of an object*, and *this is opposed to idealism*."[135] Empirical objects are outside of us insofar as they appear in space, but are not outside of us in the sense of being completely detached from our sensibility, since their spatiotemporal form is predicated on our sensibility.[136] Only the

[132] In 20th century Anglophone philosophy, a form of semantic phenomenalism associated with Russell and Carnap became popular, in which statements about objects are reducible to statements about mental states. On this version of phenomenalism, it is possible for objects to exist apart from observation, so long as statements about them are reducible to statements about mental states. See Henry Allison (2004): 38.
[133] Berkeleyan idealism entirely denies the existence of mind-independent objects. *Prol* 289-291.
[134] Karl Ameriks develops a moderate interpretation of transcendental idealism in Karl Ameriks (2012): Chapter 3.
[135] *Diss* 2:397. Italics mine. See also Kant's Refutation of Idealism, in which he establishes the perception of inner time-determination requires a permanent and mind-independent object. *KdrV* B275.
[136] Note the ambiguity in "outside of us," between that which appears outside of us and that which is a completely detached thing in itself, at *KdrV* A373.

form of the object is predicated on sensibility, however, not the object itself, and so we do not lapse into full identity phenomenalism and abandon empirical realism.

In addition to the ontological realism of the object apart from its spatiotemporal form, Kant identifies an epistemic universality to space and time that redeems him from charges of phenomenalism. Kant emphasizes that time is the universal and necessary form of all sensible phenomena, and so the axioms of time pertain as universal truths to all representation.[137] Similarly, all sensible representations fall under the geometrical principles of space.[138] The spatiotemporal world is public and intersubjective, insofar as space and time are necessary and universal features of human experience and not merely the private quirks of individuals. Space and time are not like the subjectively variable and accidental qualities of taste and color, which are not universal features of cognition but depend merely on peculiarities of individuals.[139]

In his debate with the Leibnizian Eberhard, Kant claims that Eberhard uses the term *Gründe* in a misleading way, since for Eberhard the grounds of appearances are parts of appearances, i.e. parts of empirical objects that can be discerned through the clearer lens of the understanding. The Kantian notion of *Gründe* is quite different, insofar as it is something supersensible that is not a part of empirical objects. Kant can make distinctions within empirical objects, then, between, for instance, what is subjective and objective (e.g. a pleasing taste as opposed to the shape of food) without accessing his notion of a thing in itself.[140] The supersensible status of Kant's notion of *Gründe* means

[137] See *Diss* 2:402 as well as *KdrV* A36/B52.
[138] *Diss* 2:405, as well as *KdrV* A28/B44.
[139] *KdrV* A28-9.
[140] See especially Immanuel Kant, "On A Discovery," in Allison (1973): 208-209.

that restricting cognition to appearances, and not to things in themselves, does not entail an identity phenomenalism about appearances. The raindrops underlying the rainbow is the thing in itself for Eberhard, but an objective aspect of appearances for the Kantian.[141] If there is an objective aspect within appearances, then we cannot reduce them to mental states.

The Kantian can make an appearance/reality distinction within domain of appearances: someone may have a hallucination of an alien at a certain spatiotemporal location, but the reality is that the alien is not really there, in the sense that it does not have a deeper ground within the empirical domain of possible experience. At the same time, we can posit, with our reason alone, non-spatiotemporal objects abstracted from human sensibility. Spatiotemporality has a contingency and is relative to the human subject. It does not characterize things *in general*. Things in general represent the deeper ground that is a substratum, not a part, of empirical appearances.[142]

We do not have to abandon all ontological idealism to retain empirical realism. It must be noted, moreover, that, though Melnick's view is compatible with empirical realism, it prohibits a realism about metaphysical objects. If the very idea of an object independent of the conditions of cognition is incoherent, then practical reason's attempt to supplement the causal sequence in time at the empirical level is null and void. If the idea of anything non-empirical is simply incoherent, then it would be difficult to even

[141] Ibid., 8:205, where Kant talks about a non-sensible part of empirical intuition in which it is grounded.
[142] See Rae Langton's discussion of how things in themselves do not form composite object with appearances in Chapter 9 of *Kantian Humility.*

formulate a dual standpoint theory of freedom, since the very standpoint of abstraction from cognitive conditions would be incoherent.

The judgment in this dissertation is that the notion that things in themselves are incoherent is simply wrong.[143] The thing in itself is not an object independent of all form of representation. The Transcendental Deduction does not situate objecthood as such in relation to our cognitive capacities, but instead proves the objective validity of our pure concepts in relation specifically to our sensible intuitions.[144] The procedure of the Transcendental Deduction in fact contradicts Melnick's thesis: objecthood begins as an a priori structure independent of sensible representation, not dependent on it, which is why we have to prove that our a priori notion of objecthood actually applies to sensible representation.[145] We can therefore use the category of objecthood to think of things in themselves independently of sensible representation.

In the next section, I will describe an epistemological interpretation of TI that ascribes some *intelligible* role to the thing in itself. These interpretations continue to deny that TI is a metaphysical thesis positing two separate sets of objects. But, they maintain that we can at least meaningfully deploy the notion of things in themselves, and so reject the idea that the very notion of a thing in itself is incoherent.

Henry Allison: Methodological TI. Henry Allison defends what he calls a methodological interpretation of TI. Methodological TI makes the discursivity thesis

---

[143] We see a similar view in Fichte, who thinks that a thing in itself is a thing beyond any faculty of representation. This kind of thing is not even thinkable, insofar as thinking is a form of representation. The thing in itself is therefore a "freak, a dream, a non-thought." Quoted in Karl Ameriks (2000): 180.
[144] *KdrV* A93/B125.
[145] Kant acknowledges in his famous letter to Herz that we do not derive our pure concepts from empirical objects, nor do we have archetypal intellects that creates these objects. There is therefore a problem of how our pure concepts relate to empirical objects. Guyer (1987): 23.

central to TI. The discurvity thesis implies, on the one hand, that we can consider objects as they appear, as they are given to human sensibility with its a priori forms. At the same time, we can merely think, with our understanding alone, the same object as it is independent of any relationship with sensibility.[146]

Any use of the understanding by itself to think of things independently of sensibility is merely problematic, i.e. it cannot refer to an actual object.[147] The only use of the pure understanding in its consideration of objects is to demarcate certain epistemic limits, which form the boundary line between the legitimate use of the understanding in the domain of possible experience, and its misuse as a pure understanding independently of any experience. Things in themselves are only a negative concept meant to show us what appearances are not—i.e. appearances are not detached from the epistemic conditions by which we come to know something.[148]

Allison's TI is an epistemological idealism that recognizes the reality of a priori epistemic conditions for cognition for human discursive knowers. At the same time, it rejects an ontological idealism in which the distinction between appearances and things in themselves is one between something ontologically mind-dependent, and something metaphysically absolute. Rather than defending the idea of things in themselves as some

[146] Henry Allison (2004): 16-17.
[147] Allison drew heavily from Graham Bird, who also thought that any use of the understanding to posit a thing in itself outside of experience is deceptive and problematic. Graham Bird (1962): 53. Guyer distinguishes between the positive conception of the noumenal object, and the conception of it as a merely limiting and negative. Guyer (1987): 344.
[148] This is one of four levels of meaning for things in themselves that Robert Adams identifies in Robert M. Adams (1997): 804.

sort of metaphysical supervenience base or substructure, things in themselves represent a merely formal category we reach by abstracting from epistemic conditions.[149]

TI in Allison becomes a therapeutic tool that cultivates epistemological modesty. We cannot make any metaphysical judgment about things in general—i.e. about thinghood as such, independently of how things are given to us in our particular mode of sensibility. Any sort of metaphysical judgment about thinghood as such transcends the limits of our cognitive capacity, which can only reach things given to our sensibility in space and time. Metaphysical judgments represent a God's-eye view that transcends the limited epistemic conduits available to the human discursive knower.[150]

The two-object, or ontological interpretation of the transcendental distinction, according to Allison, is a representative of the sort of overly ambitious metaphysical claims that the Critical philosophy is supposed to prohibit. To claim that there is some sort of metaphysical substructure underlying appearances amounts to reaching out beyond what is humanly cognizable. Taking the God's-eye view to posit a metaphysical substructure beyond humanly cognizable appearances, moreover, sets up a distinction between what is *really* real—i.e. the fundamental level available to an intellect unlimited by the constraints of discursivity—and what merely appears to be real. The worry is that positing this deep metaphysical structure is incompatible with Kant's empirical realism.[151]

[149] Henry Allison (2004): 17.
[150] Henry Allison (2004): 27.
[151] Henry Allison (2004): 46.

*Critique of Allison's TI***.** There is a false dichotomy here, though, that forces one to choose between a phenomenalism about appearances, or the rejection of a metaphysically fundamental level of things in themselves. One can still make an appearance/reality distinction even if one acknowledges that appearances are not metaphysically absolute. It is still possible to be able to pick out the empirical thing in itself and differentiate it from merely subjective contingencies associated with particular observers, even though we may acknowledge a more fundamental metaphysical layer underlying appearances. If the discipline of physics that characterizes the physical world is not fundamental—i.e. physicalism is not true—because there is an intelligible order beyond the contingency of spatiotemporality, this does not mean that physics is false. Falsity is not the same as having a limited domain.

It is the judgment of this dissertation, therefore, that epistemological interpretations of TI like those of Allison, which de-ontologize the transcendental distinction while at the same time giving an intelligible epistemic role to things in themselves, mischaracterize ontological TI and are insufficiently motivated. Ontological things in themselves do not automatically entail phenomenalism and the consequent loss of empirical realism. It is therefore not essential to reformulate things in themselves as merely a notion that signifies epistemic emptiness.

Moreover, securing a space for the mere thought of things in themselves, by abstracting from epistemic conditions, without making a commitment to a metaphysical relativity of the epistemic framework of the anthropocentric perspective, cannot guarantee the possibility of the real existence of things in themselves. Any kind of non-empirical causality in epistemological TI would have to be a mere conceptual perspective

rather than a genuine causal power. Under epistemological TI, we cannot be sure that an ontological space for the unconditioned is available. There can be no empirical unconditioned, since all that is in the empirical domain falls into temporal succession, and we know from the Second Analogy that temporal succession requires a causal law.[152] What is empirical cannot contain the unconditioned because appearances in the empirical domain cannot be complete, or determinate: the unconditioned in appearances can only be sought for as a problem, and not really given, and there can only be relational properties in appearances, no intrinsic non-relational properties.[153] The incompleteness of the empirical domain is established in particular in the first antinomy, in which an attempt to posit the unconditioned totality of the world in space-time leads to a contradiction between a finite and infinite world.[154] The empirical world is incomplete/indeterminate insofar as the appearances of which it is composed depend on the empirical synthesis of human understanding. The position, central to epistemological TI, that we cannot have knowledge apart from what is in space-time is consistent with the metaphysical view that space-time is metaphysically fundamental. If space-time is metaphysically fundamental, then reason's search for the unconditioned, in the form of a non-empirical cause outside of the causal series, will be frustrated.

The lack of security for a genuine non-empirical power in epistemological TI means that any kind of unconditioned capacity for agency has to be reduced to a

[152] *KdrV* A531/B559.
[153] The determinacy of things in themselves, as opposed to the indeterminacy of appearances, is discussed in Robert M. Adams (1997): 811. In Kant, the difference in terms of givenness of the unconditioned for the series of appearances and that of things in themselves can be found at *KdrV* A498/B526. The difference between appearances and things in themselves in terms of relational and non-relational properties can be found at *KdrV* A285/B341.
[154] The first antinomy begins at *KdrV* A426/B454.

perspectival, anti-realist mode in which it is asserted from a context, and is not a metaphysical fact of the matter.[155] The notion of a perspective in a metaphysical framework that lacks resources for the objective realization of that perspective raises a host of problems having to do with very fundamental and disturbing questions of what is rational and what kind of agency we really have. How is it rational, for instance, to maintain belief in a capacity that is not a metaphysical fact? Is belief in agency an illusion, like someone who believes they can fly in spite of the metaphysical reality of their earth-boundedness? Rationality, it seems, demands a purging of such illusions. The mere capacity for an agential perspective, moreover, that is illusory in light of metaphysical facts, can be possessed by an agent totally ensconced in its causal context, lacking independence as the potential ultimate source of her action. An Allisonian perspectival view on agency, therefore, fails to secure the independence of persons and threatens a Spinozism in which persons are merely finite modes of a larger monistic metaphysics.[156]

An Explanation and Critique of Metaphysical Interpretations of TI

Metaphysical interpretations represent the appearance/thing in itself distinction as involving an ontological distinction, and not merely two perspectives on the same object. Allen Wood therefore refers to metaphysical interpretations as *causality* interpretations,

[155] Allison acknowledges this at Allison (2004): 49.
[156] Ameriks notes how Schleiermacher thought that an agnosticism about things in themselves means that we should be open to the possibility of a monism about appearances and their grounds, and so Kant's philosophy is compatible with Spinozism. Ameriks (2012): 125.

because they identify appearances, entirely or in part, as pertaining to an experiential domain relative to the human subject that are caused by a deeper, non-relative metaphysical ground.[157] The transcendental distinction is not a mere epistemological distinction between the known and the unknown, but an ontological one between a fundamental level and a less fundamental, contingent level.[158]

Allais's moderate metaphysical interpretation. It is shortsighted to immediately assume that all metaphysical versions of TI are phenomenalistic. Metaphysical dual-aspect interpretations present the transcendental distinction in metaphysical terms, but without the radicalism that sets private mental entities against non-spatiotemporal mind-independent ones. Lucy Allais, for instance, thinks that Kant begins with a common-sense assumption of the existence of the mind-independent things, and that there is way these things are independent of our cognition of them. Though it is true that, for Allais, the intrinsic nature of a thing is the metaphysical ground of its appearance, this does not entail some radical metaphysical picture of phenomenalistic empirical objects and non-spatiotemporal grounds. The appearance of a thing is just its aspect in relation to human cognitive faculties, and Kant recognizes that there is a way things exist apart from human cognition.[159] The moderate interpretation Allais advances rejects the idea that thing in itself is an entirely different kind of object from its appearance, an unfathomable kind of noumenal object that does not exist in space or time.[160] At the same time, the fact that the

[157] Allen Wood (2005): 64.
[158] Karl Ameriks (2003): 69.
[159] Lucy Allais (2015): 35.
[160] Allais refers to the view that the thing in itself is a separate non-spatiotemporal object as noumenalism. Ibid., 60. Also, Rae Langton advances a metaphysical view of TI that is also moderate, in rejecting the idea that appearances have to be Berkeleyan phenomenalistic entities, entirely different from mind-independent things in themselves. Instead, appearances for Langton are qualities of things that stand in potential relation with the human subject. Rae Langton (1998): 205.

way something appears does not exactly match its intrinsic nature does not mean we have to adopt some mentalist account of appearances, such that they are intra-mental entities.[161]

Allais applies her understanding of transcendental idealism to freedom and argues that it has certain advantages over epistemological TI and radical forms of metaphysical TI. Since Allais acknowledges a metaphysical ground to the way an agent appears, she can represent freedom as a different kind of causality, rather than a mere standpoint that would be causally inert and not rationally integrated with the theoretical cognition of unbroken event-causation. At the same time, by closing the gap between the thing in itself and its appearance, so that the thing in itself is not some non-spatiotemporal noumenal entity, Allais avoid a mysterious account of agency in which a timeless will somehow produces effects in time. Allais articulates her view of free will in light of a modest account of scientific laws. They are not absolute and complete explanations of reality. Instead, scientific laws provide accounts only of subsystems that are subject to interference by external factors. It is not as though deterministic laws that account for behavior at the empirical level preclude the possibility of free will. These laws are only partial explanations of a subsystem.[162]

*Critique of Allais***.** I will put forward some objections to this moderate metaphysical interpretation of TI (moderate in the sense that it does not succumb to phenomenalism or noumenalism). First, it is not clear that Kant understands empirical

[161] Ibid., 37. In a 2004 paper, Allais also claims that phenomenalist interpretations of Kant are too idealist, whereas epistemological interpretations of him are not idealist enough. Lucy Allais (2004): 655.
[162] Ibid., 307.

laws as *et ceteris paribus* laws[163], i.e. laws that hold of a subsystem so long as all other things are equal and there is no interference outside of the subsystem. Kant seems to endorse a stronger understanding of empirical laws, which provide a complete account of the empirical aspect of the person, when he opines that, if we had complete knowledge of the influences acting on someone's power of choice, we could predict the actions of this individual with certainty.[164] If the laws governing the empirical character were merely provisional laws holding but for interference, it does not seem possible that we could predict human behavior with certainty, because there would be all sorts of hidden variables potentially upsetting our predictive capacities.

Kant states instead that there is no freedom for the empirical character.[165] The antecedently determining grounds of each action—those events that precede it in time—completely determine it. Reason, in its empirical character "is all *precisely* determined and *necessary.*"[166] Free will, then, cannot be this higher-order empirical force, able to intervene in a subsystem that, but for the intervention of free will, would follow a deterministic path. Kant thinks that, based on the empirical character, we can establish rules which serve as subjective principles of action. The empirical character, then, is not some partial subsystem subject to manipulation by an external force. If it were, it could not be the basis for explanatory rules. Kant appears to think of the empirical laws of nature as necessary: if the empirical laws characterizing natural objects are not necessary, there would be no "order of nature." The faculty of judgment presupposes, as an a priori

[163] See Pereboom's discussion of such laws in Pereboom (2001): 87.
[164] *KdrV* A550/B578.
[165] Ibid.
[166] Italics mine. *KdrV* A551/B579.

principle of purposiveness, that empirical laws are not a wild and probabilistic variety, but a systematic structure reducible to basic principles.[167] Moreover, there is evidence in the Second Analogy that Kant not only wants to establish the principle that every event has some cause, but the more ambitious principle that the same causes yield the same effects, i.e. that there are universal causal laws that project from past causal relationships to future ones. "In accordance with such a rule [governing succession] there must therefore lie in that which in general precedes an occurence the condition for a rule, in accordance with which this occurence *always* and *necessarily* follows…"[168]

Moreover, it is the case that, for Kant, any event in time is subject to a causal law. Succession requires a causal law, because time itself cannot be perceived and form a standard of orientation by which we might determine events as successive. A causal law provides the only objective basis in which events might truly be said to succeed one another.[169] If free will is just another cause in time, which intervenes in a physical system at a certain moment, it would still be subject to the causal law which makes for objective succession. If absolute freedom is to arise in time, it would have to be both a source of its own action, and so uncaused, while at the same time the fact that it objectively succeeds a preceding event requires a causal law connecting it to that previous event in a specific order. This is, of course, a paradox. But, if we are not willing to embrace a non-spatiotemporal ontology in the interests of a moderate interpretation, we end up a form of compatibilism, in which free will is in time and so subject to antecedently determining grounds.

---

[167] *KdU* 5:185.
[168] *KdrV* A193/B238.
[169] *KdrV* B234.

The fact that, given sufficient knowledge, we could predict the actions of the empirical character, indicates the empirical character is causally closed. There is no possibility of an external cause interfering in this causal series. The fact, though, that determining grounds in time fully exclude the opposites of the effects that are produced does not mean that these determining grounds are sufficient, insofar as the determining grounds by themselves cannot constitute themselves. Here we can appeal again to the indeterminacy of the empirical level. Since the whole series of conditions up to the unconditioned can only be given as a problem in appearances, we have to appeal to an intelligible, i.e. non-sensible, structure to fully account for appearances. Moreover, the fact that appearances only have relational properties shows again that they need an intelligible structure to supply a non-relational intrinsic element. So, the empirical causes in appearances might be causally closed, insofar as they exclude the opposite of what they produce in a way that is not subject to some external interference in the empirical domain. At the same time, there is a need for an intelligible ground to complete what are merely relational properties in relation to a certain observer, for a series that cannot have an unconditioned that is really given precisely because it only exists in relation to an empirical synthesis by an observer. The fact that empirical motives are never sufficient supports the Incorporation Thesis: these motives can only determine us if we incorporate them into a maxim.[170]

The shift from the causal closure at the empirical level, to the issue of sufficiency with respect to an underlying ground needed to fully constitute the empirical level, corresponds to a shift from a positivist metaphysics that asks what there is, to an

[170] Henry Allison is well-known for his Incorporation Thesis, in Henry Allison (1990): 5.

Aristotelian one that asks in what things are grounded. Positivist metaphysics rejects questions about that in which things are grounded as framework-external, i.e. questions that are outside of our conceptual framework and so meaningless. Metaphysical realism about things in themselves, on the other hand, asks questions about whether supersensible objects exist, independently of whether sentences about such objects can be true in a certain language.[171] Aristotelian metaphysics seeks first principles in the form of ontologically prior substances, i.e. something on which other things depend but which in itself is metaphysically unconditioned.[172] An Aristotelian metaphysics stresses the importance of grounding structures which may be non-empirical, and so theories can be importantly different even if they make identical commitments with respect to the empirical level subject to observation.

I put forward here a theory of free will in which the causally closed empirical character is nevertheless constituted in its being by an underlying intelligible character. I therefore appeal to a supersensible element, which is thinkable through the categories, but which transcends the empirical domain to which cognition is restricted. Introducing these dual grounding relationships—one antecedent in time, and the other a-temporal and underlying—does not entail causal overdetermination. Instead, the antecedently determining causes at the empirical level are not sufficient causes insofar as they are not part of a self-grounding series. The intelligible character acts along the lines of a Leibnizian substantial form: that is, an intelligible structure that takes the potentiality of matter and imparts to it a definite identity. The substantial form acts as a kind of control

[171] For an excellent discussion of such issues, see the Introduction, by David Manley (2009): 7.
[172] See the controversy a developed by Jonathan Schaffer (2009): 350.

mechanism of matter.[173] The idea here is that the intelligible character freely selects maxims which impart a grounding structure to the antecedent grounds at the empirical level. The intelligible character is not a mover of the empirical manifestation of the body, acting as an antecedent and extrinsic ground in time. Instead, the intelligible character acts as an intrinsic form of the empirical self, the source of its very structure and existence.[174]

The intelligible character is the possible site for the Incorporation Thesis. Maxims provide an intelligible structure in virtue of which certain belief-desire combinations produce certain actions. For example, a feeling of anger will be the antecedently determining ground of an act of violence so long as one adopts the maxim that, whenever one is wronged, one will take revenge. Kant defines a maxim as a subjective principle of volition that mediates between the a priori objective principle of practical reason, i.e. the moral law, and material a posteriori incentives.[175] The Incorporation Thesis means that the feeling of anger—i.e. a certain incentive—cannot motivate unless it is taken up by a maxim that considers the incentive in light of the moral law.

If a moderate interpretation rejects the metaphysics of non-spatiotemporality, then it cannot incorporate the account of the intelligible character I am developing. The intelligible character must be free of the causal rule of temporality in order to autonomously adjudicate between its different incentives.

---

[173] Robert M. Adams (1994): 310.
[174] The distinction between an extrinsic ground that moves something, and intrinsic ground that constitutes the being of something, corresponds to the difference between efficient and formal causes, respectively, as described in the article by John Goyette (2009): 782.
[175] *GMS* 4:401.

*Karl Ameriks: Starting with things in themselves***.** I will identify three ways in which the ontological interpretation of Karl Ameriks is distinctive. First, it begins with a common-sense metaphysical realism about external objects. Rather than beginning with epistemologically primitive items, such as impressions or representations, Kant's TI begins with judgments of ordinary experience. The thesis of ordinary experience is that there are objects outside of us in the transcendental sense that interact with the mind to produce experience.[176] The "common ground" on which philosophy begins is a rejection of more baroque metaphysical theories that deny any interaction between mind and world. For instance, Leibnizians believe in pre-established harmony, in which there is no inter-substantial interaction between finite substances. Substances can only change their own internal determinations.[177] But, Leibnizian pre-established harmony is certainly not a default position for common-sense.

Kant's idea that there is interaction between an external thing and the mind is not put forward in the form of proof, and so the worry that this position violates some epistemic restriction of the *Critical* philosophy on dogmatic metaphysics is mitigated. Ameriks states that we cannot infer, from what is empirically available to us, the metaphysical presence of a mind-independent object. Causal inferences are only valid between empirical objects linked by the causal law of the Second Analogy.[178]

The premise of mind-world interactionism would not stand up to a Humean skeptic, for instance, who builds an ontology entirely from simple impressions. The idea of a substantial form, for the Humean, is a kind of fiction posited by the imagination to

[176] Karl Ameriks (2003): 5.
[177] G.W. Leibniz, *Monadology:* Sect. 81.
[178] Karl Ameriks (2003): 26.

provide some undergirding of continuity for the impressions of an otherwise diverse successive series. Commitment to Humean empiricism leads either to a nihilism about substances—there are only events, and no underlying ontological structure—or a reductionism about them—i.e. instead of a real ontological structure, there are just bundles of impressions that appear to have a kind of continuity.[179] However, for Ameriks, Kant is modest in the sense that he is not trying to create a foundationalist system meant to eliminate even the most esoteric doubts.

The second way in which Ameriks's interpretation of Kant is distinctive is that the metaphysical starting point of a mind-independent thing means that metaphysics precedes the way in which we cognize things. The fact that our experience is generated by some pre-existing mind-independent things means that our minds are receptive and finite (i.e. there is a reality beyond them). The finitude of our perspective, and the thesis that there is a domain of substances that exists independently from, and is foreign to, our cognitive structures, leads to the notion that the spatiotemporality of our sensibility is not absolute. So, unlike Allais, Ameriks is open to acknowledging the relativity of space-time to our way of knowing. It is hard to think of space-time as a substance on a par with the metaphysical starting-point of an extra-mental thing, because it is causally inert and not directly perceptible.[180]

Kant consistently makes metaphysics prior to space-time even throughout the pre-*Critical* period. In his *New Elucidation*, he establishes the principle of succession, whereby all change, and time itself, would disappear were it not for substances in

[179] David Hume (1949): 220.
[180] Kant refers to space and time as two "non-entities" that, if we take them to be absolute realities, exist only as a formal container without any causal role. *KdrV* A39/B56.

interaction with one another.[181] In his Inaugural Dissertation, Kant claims that space and time are not something objective and real.[182] Kant identifies subreptic axioms, in which people confuse conditions for the sensible givenness of something with ontological conditions for the existence of something.[183] Transcendental idealism, for Ameriks, is particular view on the relationship of space-time to grounds. The transcendental realist about space-time embraces the view that space-time is an unconditioned and fully independent reality—i.e. something objective, an ontological condition for the existence of things. The transcendental idealist views space-time, on the contrary, as ontologically dependent.[184]

The third way in which Ameriks's version of transcendental idealism is distinctive is its arising from a regressive method. We are not to understand Kant's method as working outward from some epistemological given, towards a metaphysics based on inferences from this epistemological given. The regressive approach takes as its starting point an epistemological middle ground between a Cartesian foundationalist, who attempts to found all philosophy on the primitive *cogito*, and a starting point in sophisticated scientific judgments. Kant's starting point is the cognitions of ordinary experience.[185] Cognitions are judgments that can be true or false, and so are more sophisticated than mere representations, or bare self-consciousness.[186]

---

[181] Immanuel Kant, *New Elucidation*, 1:410.
[182] *Diss* 2:401. Kant rejects both substantivalism, or absolutism, about space-time, in which space and time exist independently of all objects, as well as relationalism, in which space-time depends on the existence of objects and expresses relational properties between them.
[183] Ibid., 2:413.
[184] Karl Ameriks (2003): 38.
[185] Ibid., 10.
[186] *KdrV* A320.

From objective judgments in cognitions[187], Kant works backwards towards the preconditions of these judgments. For example, Kant in the First Analogy begins with objective judgments of temporal relations. We not only perceive, in a subjective order, one thing after another, or at the same time. We also make objective judgments about the temporal relations in the world—e.g. we distinguish between our perception of the parts of the house, which is successive, and an objective judgment about how the parts of a house are actually simultaneous. The precondition of such objective judgments is substance that persists, against which we can measure temporal relationships.[188]

Ameriks's commitment to the regressive method as an interpretative strategy is most evident with respect to the Transcendental Deduction. As opposed to interpreters like Robert Paul Wolff[189], who view the Transcendental Deduction as a progressive move from consciousness itself to objective knowledge, Ameriks views the Transcendental Deduction as proving the applicability of the a priori concepts of the understanding to experience.[190]

The regressive method will become very important later in the dissertation when I critique contemporary libertarian accounts of free will, as invoking a neo-Humean empiricism that causes them to predicate free will entirely on events and reject a

[187] The fact that cognitions are objective does not mean they are true. *KdrV* A58/B82. The objectivity of cognition instead means they are judgments about the world, rather than mere expressions of one's feelings—e.g. I feel warmth is a sensation, whereas the idea that the sun warms the stone is an objective judgment.
[188] *KdrV* A184/B227.
[189] Robert Paul Wolff (1963).
[190] Dieter Henrich, in his canonical piece on the Transcendental Deduction, argues that the Transcendental Deduction is meant to show that there is no disproportion between consciousness and givenness, i.e. the pure concepts of the understanding apply to all appearances. The proof of the Transcendental Deduction is a two-step one that establishes first that some appearances fall under the pure concepts, then that all appearances fall under the pure concepts. No appearances can be independent of a unity in space and time, and space and time are subject to the categories. Dieter Henrich (1969): 647-9.

metaphysical substructure. A Kantian regressive approach can unpack the necessity of a metaphysical substructure underlying our cognition of events, insofar as, as we learn from the Analogies of Experience, metaphysical substructures are essential to objective time-determination.[191] One of the major concerns, with agent-causal accounts of libertarianism is that the very notion of substance-causation, in which the relata in a causal relationship are a substance and an event, is mysterious and perhaps incoherent. If a substance is to directly cause an event, it seems as though it must do so without any mechanism by which it produces the event. The exercise of the mechanism would itself constitute an event, and undermine a direct relationship with the substance.[192]

There is an implicit commitment to a Humean empiricism in this objection, insofar as it seeks to ground ontology in impressions. Substance-causation must be rejected, because there is no empirical impression with which it can be associated. The Kantian approach is different: objective judgments about the temporal determination of an event require a substance. Time itself cannot be perceived, and so only an abiding substratum of substances can serve as the standard of orientation by which we establish temporal determination as simultaneous or successive. The Kantian approach takes objective judgments as its starting point and then works backward to the necessary contextual structure for these judgments. This allows it to develop a richer ontology than an empiricism based only on discrete and instantaneous impressions. I will show how empiricism ends up with very implausible and artificial accounts of free will, insofar as

[191] The Analogies begin at *KdrV* A176/B218.
[192] Randolph Clarke tries to deal with standard objections to agent-causation in Randolph Clarke (2003): 186.

the metaphysical substructure neo-Humeanism rejects is essential for establishing agential control.

Conclusion

In this chapter, I opened with a discussion of different metaphysical frameworks for agency: an event-causal and an agent-causal one. The event-causal account leads to a crisis in which its own terms—that each event is subject to a prior event according to a causal law—mean that it cannot provide a complete account of any event. The event-causal account is also problematic insofar as it presents a trilemma of chance events, deterministic events, or an infinite regress of events, that never allows the agent to establish control. Without a metaphysical realism, in which our metaphysics is open to a supersensible ground that does not arise in time, we cannot rescue event-causation from this crisis. I then showed how different interpretations of TI try to balance, with varying degrees of success, commitments to both empirical realism and metaphysical realism in Kant. Epistemological TI emphasizes empirical realism, but does nothing, or is positively hostile towards doing anything, to secure any metaphysical realism about supersensible objects like absolute freedom. I then showed how metaphysical TI does not have to lapse into a weird phenomenalism about empirical objects and surrender any claim to empirical realism, while at the same time positing metaphysical structures that illegitimately extend the categories beyond the spatiotemporal domain. TI involves a rationalism, insofar as it articulates empirical experience in terms of pure and a priori structures of the understanding, but this is not a dogmatic rationalism that claims cognition through pure reason of things in themselves. TI involves immaterialism, insofar as the spatiotemporal

framework is not taken as an absolute determinant of things, but it is not a spiritualistic immaterialism that claims to identify a specific topography involving mental traits of this immaterial domain.[193] The limitations of theoretical reason, though, to the domain of possible experience, are significantly complemented by practical reason's access to the supersensible.

[193] Note these qualifications in Karl Ameriks (2003): 6. See the Paralogisms, for instance, in which Kant denies cognition that there is a substantial soul underlying the appearance of matter, based on the transcendental form of the "I think". *KdrV* A349.

CHAPTER 2

GROUNDING RELATIONSHIPS IN THE METAPHYSICS OF AGENCY AND EPISTEMOLOGICAL INTERPRETATIONS OF TRANSCENDENTAL IDEALISM

In the transition from the pre-Critical to the Critical period, we see a remarkable change in Kant's theory of freedom. In an early work, *New Elucidation*, Kant endorses a form of compatibilism, in which a supposedly free will is connected by antecedent grounds to the will of God.[194] In the Critical period, in both the first[195] and second *Critique*[196], Kant is committed to identifying free agency as something transcendental, i.e. outside of the causal order in time, because the spontaneous form of causality capable of self-determination cannot possibly fit into the continuous conditioned-condition causal structure in the empirical domain. The underlying shift here involves a shift in grounding relationships between the pre-Critical and Critical periods. In this chapter, I will explore the diversification of grounding relationship in Kant, which are related to the move away from a dogmatic rationalism that establishes all events in a causal series in time leading up to a necessary being, and which ultimately allows Kant to avoid the dilemma of either a compatibilist theory involving theological determinism, or an empty indifference of equilibrium in which an undetermined choice occurs by chance.[197]

---

[194] *ND* 1:397.

[195] Kant in the first *Critique* does not positively establish the existence of transcendental freedom, but shows that it is possible in the Antinomies, since the empirical series with its continuous causal structure does not exist as a metaphysical absolute. *KdrV* A558/B586.

[196] Kant's great and explicit repudiation of compatibilism comes in *KdpV* 5:97.

[197] Kant criticizes the indifference of equilibrium, in which a will is equally poised to select various alternatives and has no sufficiently determining ground directing it a certain way, as mere choice by chance. *ND* 1:403.

In section one, I will delineate the differences between Kant's early compatibilism, and the rejection of compatibilism in the Critical philosophy. In section two, I will discuss the significance of grounding relationships in relation to the shift from compatibilism to its rejection. I will point to certain virtues of the new grounding relationships associated with TI, including its ability to rescue us from a Spinozistic denial of human autonomy, an ability to break us free from the event-causal order, and also an ability to solve the puzzle in free agency of an infinite regress. In section three, I will eliminate any confusion that might conflate a Kantian appearance/thing in itself distinction, and a Leibnizian one, which also posits a non-sensible grounds of appearances. In section four, I will articulate a positive view of how the intelligible (i.e. merely thinkable) and empirical (i.e. given to human sensibility) levels in Kant's metaphysics of agency are supposed to interact. In section five, I will discuss the ability of epistemological versions of transcendental idealism to accommodate this integration of the intelligible and empirical levels.

Kant's Early Compatibilism and Its Later Rejection

*Kant's Pre-Critical compatibilism***.** We are exposed in *New Elucidation* to a kind of best-of-all-possible worlds metaphysics, in which God's originating will is subject to His perception of what is best[198], and all events flow out as necessary consequences of the original divine decision, including human action. Kant's early compatibilism

[198] This view is opposed by Crusius's voluntaristic view of God's will. See Watkins (2005): 88.

predicates freedom not on insufficient grounds that allow the human will a flexibility between different options. Instead, our freedom is based on the fact that, though our actions fall under a fixed law rooted in the divine will, they are immediately rooted in what our judgment determines to be best.[199] We see a very similar theory of freedom in Leibniz, who endorses a strict necessity whereby all events, including human action, are subject to a divine predestination. Humans, as well as God, necessarily choose what we perceive to be good. This necessity, rather than taking away from our liberty, actually bolsters it: action that is not connected necessarily to its antecedent grounds occurs by chance.[200] This pre-Critical compatibilist agent has the feature of reasons-responsiveness: that is, the fact that it acts on its perception of what is good means that it responds to reasons, and, if those reasons are different, the agent would behave differently. For some contemporary compatibilists, reasons-responsiveness is all that is needed for genuine agency.[201]

The pre-Critical picture of agency, though, is unsatisfactory, since it is scarcely distinguishable, if at all, from Spinozism. Since everything is rooted in God, and all human actions are ultimately determined by the will of God, the human agent is lost in God, becoming a mere accident inhering in the all-consuming substance of God. The very thoughts of a human being, since they are determined by the will of God, in the final

[199] *ND* 1:400.
[200] For Leibniz's articulation of this theory of freedom, see G.W. Leibniz (1956): Secs. 6-8.
[201] An agent with an irresistible physiological urge, such that, even if there were reasons to do otherwise, the agent could not help but act on the urge, would not be responsible. See Pereboom's account of Fischer and Ravizza in Pereboom (2001): 108.

analysis are the thoughts of God.[202] The pre-Critical agent has an independence of a robot programmed to respond in pre-set ways to external stimuli.[203]

The pre-Critical Kant has a dogmatic view on the nature of time, which he will reject later on as a form of a transcendental realism, i.e. an approach that attributes temporality to the absolute structure of things and not merely to human sensibility.[204] There is emerging evidence that Kant, even in his pre-Critical period, had an interest in Spinoza, and that there was a fear that the inevitable result of dogmatic metaphysics, in the vein of transcendental realism which takes the temporal order of our experience as the measure of absolute reality, is Spinozism, which subsumes agency in God and so endorses fatalism.[205] If we absolutize the empirical series in space-time, then we end up making material reality in this series, with its mechanistic necessity, a mode of extension that expresses the mode of thought. That is, transcendental realism places the empirical order on the same level with the intelligible order, since transcendental realism conflates empirical structures with thought of things as such.[206] Jacobi, in fact, claimed that one of the results of Kant's pre-Critical proof for the existence of God was Spinozism.[207] Kant makes it clear in *KdpV* that a transcendental realism about space and time leads to Spinozism: if space and time are dogmatically asserted as absolutes, then we end up with

[202] Jacobi was a critic of Spinozism, claiming that it was an unfortunate product of philosophical hyper-rationalism that denies human freedom. Michael N. Forster (2012): 59-60.
[203] See this comparison in Mele (2006): 189. For a powerful argument against compatibilism, see Pereboom (2001): 110-117.
[204] At *KdrV* A36/B52, Kant rejects the idea that time can exist independently of our intuition, as either something that inheres in things (Leibnizian time) or subsists of itself (Newtonian time).
[205] Omri Boehm (2012): 27.
[206] As Michelle Grier notes, the transcendental realist takes appearances for things in themselves, and appearances become the ultimate reality which is the only possible domain of applicability for reason's demand for the unconditioned. Michelle Grier (2001): 225.
[207] Omri Boehm (2012): 29.

having to situate the necessary being in space and time, which in turn squashes any hope for independent agency, since the actions of the agent over time will be subject to the causal control of the single necessary being.[208] If, on the other hand, we adopt TI, and treat actions over time as appearances, then God's creative power is limited to the timeless noumenal domain, where He can create agents as transcendentally free, and loses causal supremacy over the agent's actions. "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."[209]

*The rejection of compatibilism in the Critical period***.** In the Critical period, we see a staunch rejection of this early Leibnizian compatibilism. Kant claims that rooting action in the psychology of the agent, in a larger deterministic system, makes no difference with respect to freedom. If an agent is controlled by events in the past, it lacks agency, because the past is not subject to agential control.[210] It is only in virtue of transcendental freedom, the Critical Kant tells us, which has an independence from all empirical grounds in time, that an agent can be free.[211] The so-called freedom of the compatibilist agent, which is controlled by antecedent grounds and only free in terms of an absence of constraint on the agent's acting on the representation of what is best, is no

[208] *KdpV* 5:101. Spinozistic agency can "threaten our self-conception as deliberative agents…undermine reactive attitudes…there would be no reason to be moral…morality itself would be incoherent." Pereboom (2001): xviii.
[209] *KdpV* 5:101.
[210] The same premise is in Van Inwagen's consequence argument, cited in Timpe (2013): 35.
[211] For the decisive break with Leibnizian compatibilism, see *KdpV* 5:96. Even though, however, the explicit break with Leibnizian compatibilism comes in the second *Critique*, much groundwork is laid for it in the first *Critique*, and even in the semi-Critical *Inaugural Dissertation*, with its division between the sensible and intelligible worlds and the affirmation of the ideality of space and time.

better than the freedom of a turnspit—some sort of turning roasting device[212]—which moves of itself, but only insofar as it is prompted by having been wound up.[213] The idea here is the compatibilist agent has no motive power of its own, deriving it all from antecedent grounds, and so lacks any deliberative control. The turnspit agent has no independence, since her motives are subsumed in a larger, impersonal deterministic framework. A compatibilist might claim that the turnspit agent is "reasons-responsive"[214], insofar as it acts according to what it represents to be best, but a turnspit merely passively responds to reasons given to it by external processes. A turnspit does not determine itself according to certain reasons, so its freedom is only that of action, not that of the will.[215]

In order to integrate a self-determining motive power into our metaphysics, we need to establish how an event can be contingent without an antecedently determining ground that accounts for its existence. Spontaneous causality would be contingent, since its status as the source of its action would give it control over the selection of alternative possibilities which may or may not be realized. But, freedom would lose its spontaneity if it were subject to antecedent grounds. Transcendental idealism introduces a new kind of grounding relationship, in which we can introduce the self-actualizing contingency of Crusius as complementary to the order of antecedent grounds in time. Crusius thinks that

[212] Interestingly, Kant uses another device as a metaphor in the third *Critique*. Here, it is a handmill that is supposed to symbolize despotic rule. I think the handmill and the turnspit are similar in that their activity is driven entirely from external sources. *KdU* 5:352.
[213] *KdpV* 5:97. The turnspit comment marks Kant's rejection of the compatibilist ruse that a deterministic agent is not as such a compelled agent—i.e. a deterministic agent is not the same, for instance, as a drug addict compelled by an irresistible urge.
[214] Mele uses this phrase in Mele (2006): 149. The machine model of humanity was circulating at Kant's time, represented in La Mettrie's *L'Homme Machine*, and actually made its way to King Frederick's court. Cited in Rumore (2014): 83.
[215] The distinction from freedom of the will and that of action comes from Hobbes, cited in Timpe (2013): 20.

a contingent event can actualize itself without an antecedent ground, i.e. that "all opposite determinations are excluded by the mere positing of actuality alone, and, hence, that there is no need for a determining ground."[216]

The Shift in Grounding Relationships From the Pre-Critical to the Critical Periods

*Antecedent grounds in time in the Pre-Critical Kant***.** Kant articulates his metaphysics of agency in *New Elucidation* in the context of a discussion of grounding relationships. Kant defends the view that nothing occurs without a determining ground, against the view that contingent things can exist without an antecedent ground.[217] The tightly structured metaphysics Kant articulates here, in which there is no causal indeterminacy, traces the series of causally connected contingent events all the way to God, the absolutely necessary being that alone does not require an antecedent ground. This metaphysics is the dogmatic metaphysics of Leibniz, in which the logical principle of sufficient reason is taken as fundamental to the ontological structure of the world, which leads reason along the causal series all the way to a proof of the necessary being of God.[218] There can be no determinate subject-predicate relationship, according to the pre-Critical Kant, except in virtue of an antecedently determining ground. This principle, when applied to the agent, subjects it to a series of interconnected grounds, in virtue of which the agent's actions are ultimately grounded in the necessary being of God, who is

[216] *ND* 1: 397.
[217] *ND* 1:396.
[218] For Leibniz's defense of the PSR, particularly against Epicurus and his idea of a random swerve of atoms, see G.W. Leibniz (1956): Secs. 126-129.

the single being that is the ground of all possibility.[219] This structure of interconnected grounds unfolds over time, with God at the beginning of the series, because, in this early period, Kant is still a transcendental realist about time insofar as he appears to view it as an absolute structure independent of human sensibility.[220] Our actions are connected by antecedent grounds right back to the "beginning of all things"—Kant is not talking about an intelligible origin outside of time.[221]

What starts to unfold in the transition to the Critical period is the recognition by Kant that logical principles of reason need not constrain the ontological structure of the world. Dogmatic metaphysics is driven by a transcendental illusion, in which a subjective necessity recognized by our understanding is assumed to apply to the ontological structure of things (e.g. the principle of sufficient reason is the basis of an entire metaphysics of causally necessitated contingent events leading up to a necessary being). The Critical period, beginning with the *Inaugural Dissertation* and receiving clear expression in the Transcendental Aesthetic of the first *Critique*, sees a rejection of a transcendental realism about space-time.[222] The Critical Kant retains the idea that reason searches for an unconditioned, i.e. it seeks complete explanations bottoming out in something that lacks a condition, but we are shown in the Antinomies how reason's orientation to the unconditioned conflicts with the understanding's demand, in empirical cognition, to connect contingent events with antecedent grounds.[223] The fact that reason's

[219] *ND* 1:400.
[220] For an account of the development of Kant's views on space and time, see Gary Hatfield (2006): 61-93.
[221] *ND* 1:402.
[222] For the beginnings of Kant's mature view on space and time, see *Diss* Section 3.
[223] This is the dynamic of all four Antinomies. In the first, reason demands a finite universe, but the understanding cannot connect a beginning in time to an undifferentiated empty time, and cannot connect a limit in space to empty space with no object. In the second, reason demands simple parts, but the understanding cannot find a simple part in the manifold of elements that occupy every part of space. In

demand for the unconditioned conflicts with its application to the empirical domain of the understanding leads to the doctrine of transcendental idealism, in virtue of which the empirical domain is relativized to the subject, and the unconditioned can be posited in an intelligible domain that cannot be given to our sensibility, but which can only be the object of thought.[224]

Intelligible grounds in the Critical period. The shift to transcendental idealism opens up a new kind of grounding relationship that provides a theoretical framework for the move from Kant's early compatibilism to his later libertarian position. TI allows for the retention of the antecedent grounds in the early metaphysics, in which every contingent event is preceded by an antecedent ground. At the same time, though, this empirical series in time is itself grounded in a merely thinkable domain that is independent of human sensibility. This merely thinkable, or intelligible, domain, is outside of time, and so it provides for the flexibility by which one can control a contingent event without having to control its antecedent ground in the non-actual, and so uncontrollable, past. Contingent events in time have both antecedent grounds in time, as well as constitutive noumenal grounds when considered in terms of their absolute basis independent of human sensibility. The way a thing is absolutely determines asymmetrically the way it appears to a certain perspective. This new grounding relationship allows Kant to liberate himself from the limited options of dogmatic

---

the third, reason demands an uncaused spontaneous causality that completes the series of conditioned events, but the understanding demands a unity of experience in which each event has an antecedent ground. In the fourth, reason demands a necessary being as a ground of every alteration, but the understanding again demands a unity of experience that precludes a being without a condition and an infinite series of contingent things cannot have necessity. *KdrV* A426-452/B454-481.

[224] See Kant's resolution of the conflict between transcendental freedom and mechanical causality, through the diversification of the person into an empirical and intelligible character connected to different mental faculties, in *KdrV* A538/B566.

metaphysics, which, when assessing a contingent event, either assign it an antecedent ground in time linking it to a necessary being or claim it is due to chance.

Another way of articulating the added dimension to grounding relationships of the Critical period is to differentiate between a cause and a ground: while a cause is a separate event, having a temporal priority, triggering another event, a ground is a way something is (and so not entirely separate) in an absolute sense, which accounts for the way this same thing is in a certain specific circumstances.[225] Grounding, then, is not a triggering event, but a constitutive determination that accounts, in virtue of an absolute nature of a thing, for a particular manifestation of the thing, e.g. the appearance of lightning is constituted by electrons.[226] What happened is that reason's principle of the unconditioned, in the pre-Critical compatibilism, was forced into the time-series, and so could only situate agency in terms of antecedent grounds in time that form a continuous chain up the unconditioned principle of the necessary being at the beginning of the causal chain. With TI of the Critical period, we recognize the subjective-relativity of the time-series, and so reason's aspiration for the unconditioned does not have to locate the unconditioned in the time-series, but can locate it instead in an intelligible space outside of time.

Kant's new *libertarian compatibilism*[227], in the Critical philosophy, is a form of *altered-law compatibilism*, in which the lawful order in the empirical domain is shaped in

[225] We see this distinction between causation and grounding in Audi (2012): 687.
[226] Kit Fine uses the phrase constitutive determination in Kit Fine (2012): 37.
[227] Libertarian compatibilism combines elements of compatibilism, insofar as it combines the thesis that freedom and determinism are compatible, with, however, a definition of freedom that includes a robust ability to do otherwise. Wolfgang Ertl, (2014): 409. A *robust* ability to do otherwise is different from an ability to do otherwise in a compatibilist sense, e.g. the agent could have done otherwise had circumstances been different. See Mele (2006): 93 for an interpretation of robustness.

terms of its nomology by the self-actualizing choice of an agent.[228] Kant's theory is a libertarian compatibilism insofar as new grounding relationships associated with TI allow for a tight causal structure controlled by antecedent grounds in the empirical series, while at the same time positing transcendental freedom at the intelligible level. TI's grounding relationships allow, then, for the integration of a robust ability to do otherwise—a genuine power of self-determination in transcendental freedom—with a continuous deterministic causal structure in the empirical domain. Kant's theory of freedom is a form of altered-law compatibilism insofar as transcendental freedom, rather than acting as an antecedent ground in time, is actually the transcendental cause of the lawful structure of the empirical series.[229]

*The deduction of the moral law through the new notion of intelligible grounding of empirical laws***.** Kant's explicit commitment to the idea that the intelligible order grounds the empirical order, and so its laws, shows his expansion of the grounding relationships in the Critical philosophy beyond just the antecedent grounds in time. Kant writes, "…The world of understanding contains the ground of the world of sense and so too of its laws, and is therefore immediately lawgiving with respect to my will…"[230] Our membership in the intelligible world, and the fact that the intelligible world is the ground of the laws in the empirical world, is in fact crucial for the deduction of the applicability

[228] In a way, then, the way Kant articulates the relationship between the empirical and intelligible aspects of the person reflects the Leibnizian idea of a well-founded phenomenon, in which the appearance of a spatiotemporal structure is grounded on a monadic substance. However, we will see that Kant makes a crucial alteration in the Leibnizian ontology. Wolfgang Ertl (2014): 408.
[229] *KdrV* A546/B574.
[230] *GMS* 4:453. The fact that the intelligible character serves as the ground for laws of nature provides an important amendment to Van Inwagen's consequence argument, which infers, from the fact that we *cannot* control the past or the laws of nature, any product of these factors cannot be up to us. See Mele (2006): 97.

of the moral law to us. First, our participation in the intelligible world, of which we can only think about and which lacks the sensible givenness associated with the empirical world, exempts us from an illusory agency which is, in the final analysis, nothing over and above mechanistic causation and can only lie in an epiphenomenal perspective. Second, the fact that we are beings that participate in both the empirical and intelligible domains allows the moral law to take the special form of an imperative for us. If our wills were strictly intelligible, we would always follow the moral law: it would be natural to us, and not amount to an imperative. If we were strictly empirical, we would be fully subject to the mechanism of nature, and no imperatives would be applicable to a deterministic machine. But, we cannot be strictly empirical, since the empirical domain is indeterminate: it contains only relational qualities[231], and its conditioned-conditions series can only be given as a problem.[232] Since the empirical domain is structurally determined by the intelligible order, the laws of the latter pertain to the former. The moral law of the intelligible domain applies to the empirical self as an imperative, and not as a natural baseline principle, because the empirical self is drawn to go against the moral law in virtue of its sensible inclinations.[233]

With the flat pre-Critical metaphysics, which operates on a transcendental realist view of time and so tracks causal continuity only in terms of antecedent grounds in time, we can only have a compatibilist agent controlled by its past. This is the turnspit, which cannot be subject to an unconditional moral law because it is not capable of overcoming

[231] Kant states this in his Amphiboly of the Concepts of Reflection, at *KdrV* A277/B333.
[232] This fact about the empirical series is crucial to the resolution of the Antinomies. *KdrV* A499/B527.
[233] For a discussion of the deduction of the moral law through an appeal to the intelligible world, see Henry Allison (1990): 225.

its empirical motivations in the name of an absolute standard.[234] The turnspit may recognize the moral law, respect it, and take certain inclinations as conforming to the moral law[235], but the turnspit's adherence to the moral law is only accidental, i.e. a matter of luck having to do with the way the turnspit is situated in relation to the causal series leading back to the will of God. A turnspit is ultimately controlled only heteronomously, since it lines up with the moral law only insofar as it is controlled to do so by prior grounds. The way out of the turnspit indifferentism to an unconditional moral law is to recognize that the empirical level is not brutely instantiated, i.e. does not exist of itself but only in virtue of an intelligible level.[236] The limitation of the empirical order permits the existence of a transcendental freedom, which can determine itself independently of any antecedent condition. Transcendental freedom, in turn, is the necessary condition for the possibility of determination of the agent by the merely formal consideration having to do with the universalizability of a maxim, independently of any empirical incentive. Kant writes, "If no determining ground of the will other than the universal lawgiving form can serve as a law for it, such a will must be thought as altogether independent of the natural law of appearances…"[237]

What occurs in the Critical philosophy is a new kind of contingency, that allows us to recognize not only contingent relationships in time, in which a contingent event is dependent on an antecedent ground, but also to recognize that the entire time-series itself is contingent. In the Remark on the Fourth Antinomy, Kant distinguishes between a pure

[234] There is an invocation here of the "ought implies can" principle. See Pereboom (2001): 143.
[235] The epistemic mode of taking as is central to Allison's account of freedom in Allison (2003): 37.
[236] Paul Audi uses the language of brute instantiation in Paul Audi (2012): 686.
[237] *KdpV* 5:29.

sense of contingency, and empirical contingency. The former is merely abstract and independent of temporal determination: the contingent is that which has a contradictory opposite that is possible. Empirical contingency is distributed over time, so that we identify something as contingent in virtue of the fact that it requires a cause in a time prior to it.[238]

Once we adopt an idealism about space-time, such that space-time is metaphysically conditioned[239] on an underlying ground, we can introduce the unconditioned as something independent from the temporal series. The temporal series, then, would not only be contingent on the surface level, in that each of its alterations is subject to a prior cause, but the entire temporal series is contingent in that we can think of a more fundamental, and intelligible, level. This new form of contingency would not be a temporally-bound contingency, but an abstract form indicating merely that the opposite of what occurs at the spatiotemporal world is possible. Introducing an intelligible ground, which relativizes the empirical series in space-time to a mere appearance and not a metaphysical absolute, allows one to detach the empirical series from the being of God, and so avoid a substance monism in which human actions in the empirical series are modes of God subject to divine sovereignty.

It is important to understand properly the interaction between the intelligible and empirical levels, so as to avoid the confusion that heterogenous domains—one merely thinkable and unschematized, the other schematized in the spatiotemporal domain—are mysteriously supposed to interact with one another in a direct manner. Kant says that

---

[238] Ibid.

[239] See Karl Ameriks's interpretation of transcendental idealism, in which he identifies the ideal as something "fundamentally conditioned," in Karl Ameriks (2012): 84.

nature in the empirical domain resists or promotes freedom only with respect to the *appearances* of the effects of freedom in the sensible world. The pure faculty of transcendental freedom does not interact directly with the empirical domain, as something unschematized paradoxically intervening in the time-series, but is instead the ground of the appearance of the effects of freedom in space and time.[240] The diversification of contingency relationships allows us to posit transcendental freedom as a ground, instead of trying to somehow situate it linearly within the empirical causality in the time-series and so ending up in paradox.

*New grounding relationships and ontological dependence on God***.** In TI, appearances of human action are detached from God, rather than embedded in a causal series leading up to God, insofar as God's creative power generates the substances of beings, and not the way they appear. TI is therefore instrumental in challenging the implication that, if God created human beings, then our actions have an antecedent ground in God's will which is outside of our control.[241] If actions in an empirical time-series are mere appearances, they are merely secondary effects of the ontological substance of the self for which God's creative power would be responsible.[242] God's creative power at the level of things in themselves does not create a theological determinism in which our actions are part of a causal sequence ultimately traceable to the divine will. Our actions as mere appearances are not created by God, since, as Kant notes,

[240] See the note in *KdU* 5:195.
[241] See *KdpV* 5:101 for Kant's discussion of this problem.
[242] This is my interpretation of Kant's claim that it is a contradiction to assert that God is the creator of appearances—namely, that a creative power cannot cause something to exist by bringing into being derivative effects of this thing. Brewer and Watkins explain the idea that God cannot cause appearances by noting that God does not have a sensibility, as an infinite and unconditioned being, and therefore He cannot create temporal things. I fail to see the connection between inability to represent and an inability to create. Kimberly Brewer and Eric Watkins (2012): 164.

creation can only involve noumena—i.e. we do not create something by creating the way it appears. TI leaves open the possibility, then—we cannot have theoretical knowledge of it—that God does not create our appearance, but our noumenal substance, in which we are transcendentally free.[243] Kant writes, "Just as it would thus be a contradiction to say that God is a creator of appearances, so it is also a contradiction to say that as creator he is the cause of actions in the sensible world and thus of actions as appearances, even though he is the cause of the existence of the acting beings (as noumena)."[244]

*Intelligible grounds allow us to break out of the event-causal order***.** The new grounding relationships of the Critical period are powerful, insofar as the new form of contingency they introduce allows for genuine self-determination, and not merely passive subjection to antecedent grounds. In the words of contemporary free will theory, the new grounding relationships of TI allow Kant to break out of the event-causal order. Event-causation involves functions from circumstances to effects.[245] In other words, event-causation is passive, insofar as it involves sets of relationships directly connecting the internal and external circumstances of a thing, and its subsequent effects.[246] Agent-causation involves the activation of certain circumstances that empower the agent to make its own choice, and so agent-causation is not a passive transfer from circumstances to effects. In the dogmatic Spinozistic metaphysics of Kant's pre-Critical period, there is

[243] See the discussion in *KdpV* 5:102.
[244] *KdpV* 5:102.
[245] Timothy O'Connor (2000): 72.
[246] We explain why certain effects follow from certain causes in virtue of properties of the constituent objects of the causes and effects. So, a heavy branch breaks a window because of the heaviness of the branch, and the fragility of the window. Sydney Shoemaker (1980): 109. Kant makes a similar distinction, to that between agent and event-causation, in a footnote to the antithesis argument of the Fourth Antinomy. He distinguishes two senses of "begin." One is active—it produces an effect. The other is passive—the causality of the cause itself commences, or begins. In the latter case, the production of an effect is itself an event in time, subject to an antecedent ground. *KdrV* A455/B483.

only event-causation. Each determination of a state arises in virtue of the antecedently determining ground that precedes it. The antecedent ground does not merely dispose the following event to act independently. On the contrary, nothing contingent can exist without an antecedent ground that excludes whatever is opposite to the existence of the contingent thing.[247]

If we embrace a dogmatic realism about the series of events in time, the only way out of the continuity of determinism connecting events to antecedent grounds in time is an indifference of equilibrium, in which antecedent grounds do not control an agent, and so the agent is suspended in a state of indifference towards different alternatives. But, Kant claims that such an indifference would entail that the agent acts on the basis of chance.[248] The will of an agent acting independently of antecedent grounds would fall out in a way outside of the agent's control, so that, for instance, someone with a history of virtue would suddenly find herself falling into sin.

The only way out of the dilemma of Spinozism in which the agent's acts are controlled by antecedent grounds, and so the actions are merely a passive event-causation involving a transfer of causes to effects based on properties, and actions driven by mere chance, is the metaphysics of transcendental idealism. TI moves beyond the flat metaphysics of the pre-Critical period, and introduces an ordered metaphysics in which the time-series is a merely relative phenomenon, leaving open a domain of non-empirical features that allow the human person to be a hybrid cross-contextual being.[249] The

[247] *ND* 1:396.
[248] *ND* 1:402.
[249] Karl Ameriks identifies at least four levels of being opened up by transcendental idealism: the most fundamental level of intrinsic/noumenal being, the relational nonempirical features of transcendent

relationships between beliefs and desires, and actions, which are temporally determined and occur at the empirical level, is then contingent on an intelligible level that is not in time. Insofar as this intelligible level is not in time, it is not subject to the causal law necessary for objective succession in time. The intelligible level can therefore be the site of genuinely self-directed agent-causation.

*Intelligible grounds and the infinite regress problem of Galen Strawson***.** The diversification of contingency allows the free will theorist to address both concerns with chance, as well as the concern of an infinite regress that Galen Strawson identifies.[250] At the empirical level, we retain the pre-Critical principle that antecedent grounds determine action, and so actions do not fall out according to chance. As for the infinite regress, Strawson establishes that genuinely free action must be rational. In order for action to be rational, Strawson demands that action stem from the way one is mentally. But, in order to be rationally in charge of this mental content, it must also stem from the way one is mentally—it cannot stem, for instance, from some physiological drive. And this second-order mental content must not be out of one's control, if one is to be truly free, etc. This argument leads to a denial of freedom, because it becomes absurd to attribute control of an agent over ever more remote preconditions of her mental states.[251]

The infinite regress problem is generated by presupposing an empirical understanding of the conditioned-condition relationship. The mental content is understood in the form of antecedent grounds. Strawson's principle is that nothing can be

beings, such as God's will, non-empirical features of hybrid empirical-intelligible beings (human persons), and finally, the most relativized level of empirical reality. Karl Ameriks (2012): 243.

[250] Galen Strawson (1994): 5-24.

[251] Strawson is looking for "heaven-and-hell type moral responsibility," not a compatibilist kind, that would justify the possible damnation or eternal reward of an individual. Quoted in Mele (2006): 131.

the cause of itself, and so, for the will to act, it must have some external prompting. If one is to be truly morally responsible, one cannot be subject to external causation, and so one must have control over the very springs of one's will. However, this is the empirical notion of contingency, in virtue of which something is contingent in virtue of having an antecedent ground in time. The intelligible notion of contingency, on the other hand, does not have to do with causal relationships over time. Instead, intelligible contingency means that the contradictory opposite is possible.[252] The intelligible level does not operate in terms of a temporal causal sequence. We can establish rational control through an intelligible ground, which would, since it is outside of time, not require an implausible reaching back in time, since its contingency is an intelligible one that does not require antecedent grounds. So, the agent, in TI, can rationally structure its actions, not as an infinite series of antecedent grounds reaching back in the event-sequence, but as an intelligible form that underlies the event-sequence and so does not need antecedent grounds. The intelligible ground acts roughly like the criterial causation of neuroscientist Peter Tse: it is not a transfer of energy over time in a sequence of events, but instead an informational recoding that sets criterial parameters for the event sequence.[253] The agent is not a *causa sui* but exercises oversight over the event-sequences that characterize it.

Having developed an account of the new grounding relationships introduced by TI, I need to make a careful distinction between the Kantian and the Leibnizian transcendental distinction. In this vein, it is worth noting that Kant engaged in a polemic with a Leibnizian by the name of Eberhard. Eberhard's core thesis in his debate with

[252] *KdrV* A460/B488.
[253] Tse (2013): 38. Tse solves Strawson's impossibilist argument at 152.

Kant was that Kant's transcendental distinction was an unoriginal appropriation of the appearance/thing in itself distinction already evident in Leibniz. The Leibnizian orthodoxy is that the sensible world is a composite, the fundamental parts of which are intelligible monads, i.e. mental substances which have only powers of representation independent of any relational qualities involving a physical place or shape.[254] Eberhard argues that Leibniz, just like Kant, believed in non-sensible grounds of appearances.[255] Yet, I have already shown how Kant's early compatibilism, which he fiercely rejected in his Critical period, mirrored the compatibilism of Leibniz. In *KdpV*, Kant refers to Leibniz *spirituale*—i.e. monads—as subject to the "mechanism of nature" and as having the so-called freedom of a mere turnspit, since Leibniz does not qualify the necessity of nature in time as a mere appearance.[256] It is therefore necessary to differentiate Kant's distinction between sensible appearances and non-sensible grounds from that of Leibniz.

Differentiating Between Leibnizian and Kantian Non-Sensible Grounds

At the end of the Transcendental Aesthetic, Kant engages in a ringing rejection of the "Leibnizian-Wolffian" philosophy with respect to the difference between the sensible and intellectual faculties.[257] Kant establishes both an epistemological and a metaphysical severance between appearances and things in themselves. No amount of scrutiny of an appearance can give us any knowledge of things in themselves (epistemological

[254] See Kant's discussion of monads at *KdrV* A274/B330.
[255] See Rae Langton's discussion of this controversy in Rae Langton (1998): 191.
[256] *KdpV* 5:97.
[257] *KdrV* A44/B61.

severance). If we were to remove our subject, then all relations of objects in space and time, and space and time themselves, would disappear (metaphysical severance).[258]

The Leibnizian-Wolffian philosophy, on the other hand, seems to posit a metaphysical and epistemological identity between things in themselves and appearances. The sensible appearance of a thing in itself is merely a confused version of what it really is in itself.[259] Kant compares the Leibnizian appearance-thing in itself distinction to the difference in the concept of right between a healthy understanding and subtle speculation.[260]

We see the same polemic in Kant's engagement with the Leibnizian Eberhard. The ground of sensible representations for Eberhard is given in these sensible representations. Things in themselves, though fundamental to representations, are internal to them. For Kant, on the other hand, things in themselves represent a supersensible ground of sensible representations that are not internal to them.[261] Eberhard likens the distinction between an appearance and its non-sensible ground to the relationship between a crowd of men viewed from a distance, in which individuals are blurred, and the same crowd seen from up close.[262] The appearance of the crowd, and the ground of this appearance, in the form of articulated individuals, are also metaphysically identical—the only difference is one of clarity. But, Kant makes it clear that his things in themselves are not empirical things in themselves. A set of raindrops appears as a rainbow in certain

[258] *KdrV* A42/B59. We will see later, in chapter three, how Kant relies on the first antinomy, involving the magnitude and duration of the world, to definitively debunk the idea of the neglected alternative, i.e. that space and time are metaphysically unconditioned features of reality in itself.
[259] *KdrV* A43/B60.
[260] *KdrV* B61/A44.
[261] Allison (1973): 215.
[262] Cited in Rae Langton (1998): 191.

atmospheric conditions, and so in a way is a ground of these appearances. But, raindrops themselves are appearances, insofar as they have shape and occupy space.[263] Things in themselves for the Kantian are not parts of appearances, derivable through a clearer scrutiny of them, but transcendent grounds that are undetectable in principle through the medium of spatiotemporal experience.

In the Amphiboly of the Concepts of Reflection, Kant makes his differences with Leibniz clearer. Leibniz established the principle of the identity of indiscernibles because he takes objects to be intelligibilia, i.e. objects of the pure understanding, and does not differentiate them with respect to their location in space-time. The appearance of an object for Leibniz is just a confused version of the object, and so does not possess any kind of individuating capacity by which we can distinguish between two different objects. For the Leibnizian, relational qualities of an object in space-time are reducible, so that they provide no basis for differentiation over and above the intrinsic and abstract nature of the thing. For the Kantian, relational qualities having to do with a substance's situation in space-time are irreducible, and so they constitute a way to differentiate objects independently of their abstract and intrinsic nature.[264] For Leibniz, then, there cannot be two separate identical objects, since mere spatiotemporal location is totally reducible to the intrinsic nature of objects, and so provides no real basis for differentiation. Kantian objects, though identical in terms of their abstract nature, can be

[263] *KdrV* A46/B63.

[264] See the discussion on the contrast between Leibniz and Kant on the reducibility of relations in Rae Langton (1998): 108.

differentiated according to their location in space-time, since space-time represents a set of relational qualities that are irreducible to their non-sensible grounds.[265]

The key difference between Leibniz and Kant has to do with the relationship between matter and form as dual forces of the constitution of objects. Matter is the determinable—stuff that is in itself shapeless—and form represents what determines—i.e. imparts shape.[266] For the Leibnizian rationalist, the understanding does not have a merely immanent use within experience, but instead makes immediate contact with things. The unconditioned reality of simple monads, with a power of representation, is the matter with which the understanding initially makes contact.[267] The form of space-time arises from the simple substances as the structure of their outer relations and the community of their inner states.[268] Intelligible substance is the matter, then, which precedes, for the Leibnizian rationalist, the form of space-time, which arises from relationships between and within substances.

For the Kantian, though, the relationship is reversed, since form precedes matter. The Kantian does not believe in an immediate cognition of things through the understanding. Instead, the Kantian insists that we cognize only things that appear within the space-time structure of sensibility. The form of space-time, as the structure of sensibility, precedes the sensations that reach our sensibility. The spatiotemporal form preceding matter is abhorrent to the rationalist, who thinks that we immediately grasp

[265] *KdrV* A264/B320.
[266] *KdrV* A267/B323.
[267] We can see this confidence in reason's ability to directly track ultimate reality in the rationalist Eberhard, who insisted that a principle of reason, such as that of contradiction, though it is merely logical, also applies to real grounds even without a corresponding sensible intuition. Allison (1973): 195.
[268] G.W. Leibniz (1956): Sec. 48.

things with our intellect and space-time is merely a confused version of monads grasped by sensibility. Sensibility is just a downgraded form of the understanding for the rationalist, insofar as it grasps the thing in itself, just less clearly. For the Kantian, the form of space-time makes possible the appearance of the sensation, i.e. matter. The spatiotemporal form, again, is irreducible to the intrinsic nature of the object, since the spatiotemporal form is imparted by human sensibility to constitute a separate appearance.[269]

The difference in reducibility for spatiotemporal relations has a profound impact on our metaphysics of agency. Space-time for the Leibnizian, as fully reducible to the intrinsic nature of things, is not an individuating or constitutive property. It is merely empty property that provides an epistemic window—albeit clouded—on to things in themselves. If space-time does not individuate an ontologically distinct property, then we are not able to separate space-time from the thing in itself in a way that allows for the possibility of an a-temporal cause at the intelligible level not subject to the causal law. If space-time represents a confused version of the thing in itself—i.e. a sort of blurring, as when we observe a crowd from a distance—and not a separately constituted property, we cannot establish the non-sensible ground of the spatiotemporal appearance as itself a supersensible ground outside of space and time. As Kant notes in his polemic against Eberhard, the thing in itself for the Leibnizians is within the sensible intuition, as a fundamental part attained through the clearer perception of reason. Eberhard wished to

[269] The irreducibility of the form of space-time to the intrinsic nature of objects comes from Kant's early metaphysics, in which he establishes that the existence of a solitary substance does not entail relationships with any other substances. For the Leibnizinan, the relational qualities of substances are reducible, and entailed by, their intrinsic and solitary nature. See the discussion in Rae Langton (1998): 109.

maintain an ability of reason to have cognition of certain rational concepts—e.g. what is simple—by situating these rational concepts within sensible intuitions, as more fundamental parts. But, Kant rejects this assimilation of rational ideas to the sensible intuition, claiming instead that any part of a sensible intuition, even its fundamental components, cannot be a supersensible idea of reason. "If the whole is an object of the senses, all of its parts must necessarily be so likewise."[270]

A distorted conception of the Kantian relationship between appearances and things in themselves circulated shortly after the Critical philosophy emerged. This distorted conception held that the appearance was the form of the thing in itself, which was represented as the matter of the appearance. So, the spatiotemporal appearance represents the structure of the raw stuff in things in themselves.[271] Clearly, though, this is similar to the Leibnizian transcendental distinction we encountered in the Amphiboly, in which things in themselves are primary and then take on a spatiotemporal structure in virtue of their pre-existing relationships.[272] Leibnizian space-time is an ontologically internal structure of things in themselves, lacking in independence, and merely a degraded representation of what the intellect grasps clearly. The distorted conception of spatiotemporal appearances as the mere form of things in themselves led to the thought that Kant's transcendental idealism could be assimilated to a Spinozistic monism in which God and Nature are two aspects of the same entity.[273]

---

[270] Allison (1973): 205/pg. 122.
[271] Ameriks discusses this distorted conception in Ameriks (2012a): 125.
[272] *KdrV* A276/B332.
[273] Reinhold perpetrated the falsification of the appearance-thing in itself relationship as a form-matter relationship, and Schleiermacher spread the idea that Kant could be interpreted as a Spinozist. Karl Ameriks (2012): 47.

For the Kantian, though, space-time is the form of appearances, not things in themselves. It is not merely a degraded perspective on things in themselves, but a constitutive form of an appearance, which is distinct thing in its own right. This separate constitution of the spatiotemporal appearances is very important in differentiating Kant's transcendental distinction from that of both Leibniz[274], as well as Spinoza. In the latter two philosophies, the transcendental distinction involves a merely perspectival shift on the same thing—in Spinoza, the appearance of extension is just a parallel reiteration of non-extended thought, so that God and nature are fundamentally interchangeable.[275]

What seems to happen with this degraded role for space-time in Leibnizian phenomenalism is that it flattens the metaphysical picture. The temporal series, rather than being a distinct metaphysical track with a constitutive role, is actually transparent with respect to the monad. Kant needs to separate temporality from the ultimate structure of things in themselves, in order to avoid positing the unconditioned within the temporal series. Kant is not alone in attributing to Leibniz the idea that temporality is fundamental, and characterizes the mental states of monads, and not just their confused perceptions. Intra-monadic time is, in fact, the standard interpretation of Leibniz.[276]

The transition from a one-to-one mapping relationship of the empirical time-series on to things in themselves, to a Kantian grounding relationship between a distinct supersensible realm and empirical objects, marks a transition from a rationalist

[274] We have to recognize that we are dealing with Kant's interpretation of Leibniz, which may be tendentious and based on partial sources. Anja Jauernig speculates that Kant, in his polemical remarks against Leibniz, is not attacking Leibniz himself, but instead the Leibnizian-Wolffian *Schulphilosophie*. Anja Jauernig (2008): 50.
[275] Benedictus Spinoza (1985):Part IV, Preface.
[276] Eric Watkins (2005): 126.

cosmology in which we trace individual events back to their origination in God, to a noumenalist cosmology in which any kind of theological basis of reality is relegated to an un-cognizable supersensible domain. We see, accordingly, a best-of-all-possible worlds metaphysics in the pre-Critical Kant, in which even evil events are rooted in a strict causal structure leading to God, to a Critical Kant who denies that cosmic justice is something of which we have theoretical cognition.[277]

Positive Conception of the Integration Between the Empirical and Intelligible Levels in Kant

The introduction of an intelligible contingency, by which we can diversify away from the antecedent grounds in time in the pre-Critical Kant, and an important distinction from Leibniz which establishes the non-sensible ground as ontologically external to the spatiotemporal appearance, requires in this section a positive account of how these different contingency relationships might be integrated into a metaphysics of agency. I will identify and describe the following characteristics of this positive account: 1) An asymmetrical determinacy/indeterminacy relationship between things in themselves and appearances, and 2) the role of maxims as an activity, in the manner of a substantial form, that shapes the relationships between events at the empirical level. I will close this section with a transition into the fifth and final section, by noting that an asymmetrical

[277] Immanuel Kant, *ND* 1:404. For a description of how the Kantian grounding relationships between appearances and things in themselves differs from the merely logical Leibnizian one, see Anja Jauernig (2008): 45.

metaphysics involving intelligible maxims[278] acting as substantial forms is not compatible with a Newtonian mechanistic worldview that banishes substantial forms. Interpretations of TI that cannot banish the Newtonian worldview therefore cripple our ability to adopt a metaphysics of agency involving substantial forms.

*Asymmetry between appearances and things in themselves***.** The intelligible level is integrated with the empirical level in an asymmetrical relationship, in which appearances at the empirical level are contingent and specified features of a more fundamental and absolute intelligible level. Things in themselves represent the synthesis of the pure understanding, by which we think of things in general apart from any conception to sensibility.[279] and so have a determinate status in virtue of existing independently of a relation to sensibility. Appearances are intrinsically indeterminate, insofar as they are realized in the application of categories to the forms of a specific kind of intuition.[280] It makes sense that the nature of the intelligible level, insofar as it represents reality independent from all sensibility, would form the basis of the way the empirical level appears—i.e. the empirical level appears one way, and not the other, in virtue of the nature of its intelligible basis. An example of the asymmetrical grounding relationship between things in themselves and appearances comes in Kant's argument against the idea that his empirical bodies are illusions. One way in which he debunks this idea by noting that the predicates of an appearance can be ascribed to a thing in itself, as

[278] Maxims are structured hierarchically, with different levels of overlap between the intelligible and empirical characters. In the *Religion*, Kant mentions a noumenal deed outside of time that is the ground of every use of freedom in experience: 6:22. This root of all uses of freedom disposes us to privilege either the moral law or our inclinations. The subjective character in turn underlies empirical manifestations of particular maxims in time—e.g. a professor vindictively grades a paper, and so empirically manifests the subjective root of her empirical choices to be evil.
[279] The intellectual synthesis is found in the B-Deduction at *KdrV* B151.
[280] Eric Watkins (2005): 314.

long we are discussing the way the object appears to our sensibility. We cannot ascribe these predicates as direct properties of the object in itself, just to the way it appears to us. Nevertheless, the fact that we can ascribe redness to the thing in itself means that redness is not a mere phenomenalistic idea, but is grounded in things in themselves.[281] In the asymmetrical relationship between appearance and things in themselves, Kant is showing a Leibnizian heritage, insofar this asymmetrical relationship resembles Leibniz's view of the correspondence between bodies and monads. The asymmetrical relationship in the Leibnizian body-monad distinction is expressed in the fact that the appearance of bodies supervene on monads, and not vice versa.[282] Kantian asymmetry of appearances and things in themselves, however, is importantly different, insofar as there is a real grounding relationship between them, and not a merely harmonious correspondence. Kant complained that monadic substances, being merely capable of representations, lack any force by which they can form real causal connections with other substances.[283] It is also important to note that, even though Kantian appearances are grounded in things in themselves, we cannot establish an isomorphism between them, such that we are able to have specific and detailed theoretical knowledge of things in themselves through their appearances. The idea that an appearance can indicate to us that a thing in itself is appearing is a minimal kind of knowledge—a mere existence claim—that doesn't imply any sort of spiritualistic immaterialism claiming to have knowledge, for instance, of the

[281] *KdrV* B70.
[282] See the discussion of this Leibnizian heritage in Rae Langton (1998): 80.
[283] Kant also mentions this in the Amphiboly section of the first *Critique*, *KdrV* A274/B330. Anja Jauernig (2008): 49.
[283] *Diss* 2:407.

intrinsic nature of the self.[284] As Kant says in the B-Preface, if we cannot think of things in themselves as underlying appearances, there "would follow the absurd proposition that there is an appearance without anything that appears."[285] In the issue of free will, in particular, there is a close relationship established between appearances and things in themselves, since the appearance of action of a person on the empirical level is supposed to be a manifestation of the same person at the intelligible level. This kind of isomorphism, though, can be established through practical reason's cognition of its own freedom.

Unless free will is to be merely epiphenomenal, and not exert any causal influence on the empirical order, it must be in a grounding relationship with the empirical level. A metaphysics of causal isolation would entail a free will that freely chooses its maxims independently from the causal law of time, but which can only watch helplessly as its empirical character follows a parallel course outside of its control.

*Maxims as an activity and a substantial form***.** We know that Kant's theory of physical influx—i.e. the idea that finite substances enter into real causal interaction and not mere harmonious correspondence[286]—was importantly modified from that of some of his predecessors, insofar as Kant thought that reciprocal dependence between substances does not arise from the mere existence of substances alone. Reciprocal dependence is an activity.[287] Maxims chosen in light of the intelligible character acting under the idea of freedom—e.g. always take proportionate revenge for slights incurred—can be understood

[284] Ameriks rebuts Aquila's claim that analysis of appearances cannot lead to any knowledge of things in themselves, even their mere existence, in Karl Ameriks (2003): 84.
[285] *KdrV* Bxxvi.
[286] Eric Watkins (2005): 156.
[287] *Diss* 2:407.

as a kind activity that shapes relationships between desires-beliefs at the empirical level and actions—e.g. the event of wounded pride leads to the event of retaliation. What is chosen in light of the intelligible character is not causally inert but shapes causal structures at the empirical level, since the Incorporation Thesis requires that empirical impulses cannot directly determine action, but must be mediated through a maxim.[288] We can turn to Leibniz's idea of a substantial form to develop the idea of the activity—not the mere existence—in virtue of which maxims at the intelligible level ground the empirical level. Substantial forms are part of the heritage of the Aristotelian-Scholastic tradition, in virtue of which a form acts as a cause in imparting an intelligible structure to matter. Form actuates matter by imparting an identity to an otherwise formless stuff that is merely a potential thing. The substantial form acts as a kind of control mechanism, guiding the development of a thing according to a certain core identity.[289] For Leibniz, a true unity of a thing does not arise merely from spatial contiguity, e.g. two diamonds do not form one substance in virtue of being placed into one ring. It is a substantial form underlying these different parts that provide the unification that makes them pertain to one substance.[290] Leibniz, in fact, distanced himself from the mechanistic philosophy of Descartes, in which the original impetus for motion, as well as the laws of motion, are sustained by the continual activity of God. Leibniz added to this picture of inert bodies an underlying force, immanent to individual substances, in virtue of which they constitute the ground of their own predicates. This unifying force gives individual bodies an

[288] Maxims are a rich topic of study in Kant's practical philosophy to which I will return later. Henry Allison defines a maxim as a self-imposed rule of action, linking situation-types to action-types. Henry Allison (1990): 89-90.
[289] Robert M. Adams (1994) 310-11.
[290] Daniel Garber (1985): 35.

independence from theological determination by God's will, and so saves Cartesian mechanism from a Spinozism in which bodies and their activity are, in the final analysis, attributes of God.[291]

We can see Kant's sympathy for the idea of substantial forms, and the related explanatory insufficiency of merely extended bodies, from his earliest works.[292] In these works, Kant establishes that forces uniting the world are prior to extension, and so he inverts the Humean strategy of beginning ontology from spatiotemporal particulars.[293] Kant sides with the Leibnizians in the famous *vis viva* debate, in which two competing natural philosophies were juxtaposed: the Leibnizians favored an account in which what is ontologically primitive is not extension but a force which brings unity to a substance, whereas Cartesians favored an account in which the fundamental determination of all configuration of matter is extension. For Kant, force is to be understood as an activity that is more fundamental than motion, and this allowed him to bridge the gap between the non-extended mind and the body. If force were mere motion, there would be no way to account for how external bodies can create unextended ideas in the mind.[294] Kant goes on to establish that, without a force in virtue of which substances stand in a relationship of mutual interaction, there can be no external relations, and so no space.[295] Matter for Kant is constituted by forces: an attractive force keeps matter from dispersing, and a repulsive force keeps matter from contracting to an infinitesimal point.[296]

---

[291] Ibid., 76.
[292] *IK* I:7.
[293] Ibid., I:17.
[294] See the introduction by Eric Watkins in *IK* 5, as well sections 5-7 in chapter one.
[295] Ibid., I:23. See the discussion of *IK* in Eric Watkins (2005): 104.
[296] See Rae Langton (1998): 39.

I think the idea of intelligible maxims shaping the empirical order can be understood as an extension of the idea of forces unifying extended bodies in Kant's early scientific explorations. Maxims are subjective principles an individual imposes on himself, and are to be distinguished from the moral law, which is the objective normative principle pertaining to all rational beings.[297] The moral law pertains to the will in its legislative function—i.e. *Wille*, which is that aspect of practical reason that prescribes the categorical and hypothetical imperatives—while it is the executive function of the will—*Willkür*—that selects maxims on which to act which have varying degrees of conformity to the moral law.[298] In recent commentary, maxims are portrayed as expressing relationships between types of situations and types of action, rather than particular intentions. Onora O'Neil, for instance, identifies maxims with underlying intentions that guide individual and more specific intentions.[299] This species-genera categorization of maxims corresponds to the root, or radical, maxim Kant talks about in his *Religion*, and the specific empirical manifestations of this root. The radical maxim is a fundamental choice of character, that applies universally to each individual use of freedom.[300] Though maxims are more or less amenable to specific empirical manifestation, all involve the intelligible character, since they subject empirical incentives to evaluation in light of the moral law, and so presuppose the idea of freedom.[301]

---

[297] *GMS* 4:421n.
[298] See the discussion in Henry Allison (190): 129. Maxims represent the subjective principles of the agent, whereas the moral law of *Wille* represent an objective principle. Allison (2008): 450.
[299] Onora O'Neil (1983): 394.
[300] *R* 6:25.
[301] The freely acting being, who incorporates empirical incentives into her maxims, puts herself in an order of things outside of nature, i.e. an intelligible order. *GMS* 4:457.

The idea that maxims are underlying general rules, to which multiple individual intentions conform, fits nicely with the metaphysics of agency I am developing. A maxim can serve as the underlying structure, associated with the agent qua substance and so temporally indeterminate, that partially accounts for individual intentions that arise in time. While maxims are general, and in fact participate in hierarchical relationships akin to genera and species, they are not unchangeable *Lebensregel* (i.e. life-rules), and so are subject to the free choice of the *Willkür*.

Maxims, in the form of a noumenal choice of subjective character that is the root of all decisions, as well as maxims falling under the general maxim that mediate between the intelligible character's grasp of the moral law and specific empirical incentives, add an essential dimension to empirical relationships. Certain empirical relationships occur in one person, and not another, not just in virtue of the happenstance of empirical causation, but in virtue of chosen maxims. This added dimension allows for a genuine metaphysics of agency: if people are just subject to empirical causes, their actions happen to them, either as inevitable products of deterministic series or as chance products of indeterministic series. Maxims can ensure that the empirical series leading to actions belong to agents.

Of course, this sort of grounding relationship is not compatible with certain interpretations of transcendental idealism. If transcendental idealism is a merely epistemological doctrine that tracks the basic conditions of knowledge, and articulates two epistemic viewpoints in relation to these conditions, then no grounding relationship between the intelligible and empirical levels is forthcoming. Epistemological interpretations have an aversion to metaphysics we see in early 20th century movements

like logical positivism, which see metaphysics as just an idle game divorced from the conditions of meaningful utterance, i.e. empirical data and pure analyticity.[302] I will transition into a discussion of the limitations of the methodological interpretation by way of controversy in metaphysics over the existence of substantial forms. Isaac Newton rejected this existence as unintelligible, and claimed that we could establish individual substances in virtue merely of extension in space and the divine will. There need not be any intelligible character structuring otherwise formless matter, since a body can exist purely in virtue of extension.[303] In Newton, there emerges a worldview in which space and time are absolute and uncreated realities that are essential modes of the existence of God, and so form the fundamental fabric of being.[304] If the Newtonian worldview makes extension and its configurations absolute, by making space-time necessary modes of any existence, then it makes no sense to claim that there are metaphysical depths beyond motion in space-time.

The Newtonian view is therefore a threat to Kant's theory of freedom, insofar as this theory incorporates a non-empirical cause, acting as a substantial form, and downgrades the empirical level of space-time to a metaphysically conditioned status. A form of transcendental idealism that cannot reject the Newtonian view is therefore problematic. Space and time are not modes of existence as such, but conditions of human sensibility, for Kant.[305]

[302] A discussion of the evolution of this anti-metaphysical movement can be found in Huw Price (1997): 4.
[303] Isaac Newton (2004): 29.
[304] Andrew Janiak (2008): 147-8.
[305] See Kant's discussion of the tortuous and, on his view, unnecessary debates about the location of immaterial substances, given the absolutization of space-time. *Diss* 2:414.

Epistemological TI and the Metaphysics of Agency

We need to evaluate epistemological TI in light of the metaphysics of agency I have been discussing. I have identified a crisis in Kant's understanding of freedom, involving a rejection of the compatibilism, that is related to the profound break he had in his Critical period with a Leibnizian transcendental distinction. The question in this section will be to what extent the epistemological interpretation of TI can capture these revolutionary developments in Kant's fundamental commitments. I will evaluate epistemological TI on the following issues: 1) Practical reason's cognition of itself as absolutely free, and 2) the ability to deduce the unconditional moral law as apodictically certain.

TI is supposed to deliver reason from a situation in which it is only able to posit the unconditioned within the series of temporal causality. This is what happens in the pre-Critical Kant: reason's demand for an antecedent ground for every contingent event leads to the situation of human action in a causal series stretching back, and rooted in, the will of the necessary being. Remaining dogmatic about temporal relations forces us to articulate reason's demand for the unconditioned in terms of conditioned-condition relationships in time, and this leads to a kind of Spinozism, in which the agent is subsumed in the necessary emanations of a single substance.[306] TI is also supposed to deliver us from Leibnizian intellectualism, by which space-time is just a degraded version of things in themselves, which does not play a role in constituting a distinct appearance.

[306] *KdrV* A460/B488.

Epistemological TI claims that space and time in Kant are recast as merely epistemic conditions, rather than having a distinct ontological status. Instead of establishing space and time as metaphysically conditioned realities, epistemological TI avoids the game of ontology altogether.[307] The notion that space and time are metaphysically conditioned is actually included within the transcendentally realistic project of establishing ontological truths that seek a God's eye view.[308]

Epistemological TI in this respect captures the critique of traditional ontology that is definitely a part of the Critical philosophy, i.e. it rejects attempts to gain metaphysical cognition of things as such through a mere rationalist analysis of the concept of thinghood.[309] At the same time, the epistemological modesty of methodological TI means that it may be the case that the spatiotemporal domain is an unconditioned reality. There is no incompatibility in making the epistemological claim that we can only know what is in space and time, and accepting the thesis that space and time are metaphysically unconditioned. The epistemological theorist would object that establishing the metaphysically unconditioned status of space and time represents a God's eye version of ontology. However, the agnosticism about the metaphysical status of space and time fails to prove the possibility of an intelligible space as genuinely metaphysical, in which reason can posit the unconditioned as a real object and avoid Spinozism. Failure to prove the possibility of this intelligible space has important implications for the epistemic status of the unconditioned and so the integrity of reason.

[307] Henry Allison (2004): 98. See also Henry Allison (2006): 7.
[308] Henry Allison, *KTI*, 98.
[309] Kant says that the once-proud discipline of ontology must be replaced with an analytic of the pure understanding at *KdrV* A247/B304.

Epistemological TI and practical reason's cognition of itself as absolutely free.

Kant claims that absolute freedom, in light of practical reason's subjection to the moral law, is a fact, on the same epistemic level as what is based in intuition and so cognizable for theoretical reason, i.e. a priori geometrical truths as well as empirical cognitions in experience.[310] We have cognition of our freedom because we have an intuition of the moral law, which is the *ratio cognoscendi* of freedom.[311] Other supersensible objects, such God and the immortality of the soul, are accepted as inferred[312] components of a rational faith based on reason's cognition of itself as absolutely free. These postulates are commanded by practical reason.[313] The epistemic status of absolute freedom, in addition to the *Vernunftglaube* based on it, is very different from, for instance, an a priori opinion. An a priori opinion is an opinion based exclusively on speculation, about a matter that is not a possible object of experience.[314] Since an a priori opinion does not have a tautological validation by definition, and cannot have any basis in experience, it is a mere arbitrary invention.[315] The cognition of oneself as free, moreover, is very different from a persuasion from the first *Critique*, which is a merely subjectively valid belief that has no objective grounds that command universal assent.[316] Nor is absolute freedom a mere hypothesis: it is based on the apodictically certain moral law, whereas a hypothesis, in the

[310] *KdU* 5:468. Lara Ostaric charts the evolution in Kant from the Canon to the third Critique, in which Kant intensifies his commitment to the veridicality and objectivity of practical cognition. Ostaric (2017): 101.
[311] *KdpV* 5:4, note.
[312] God and immortality are inferred, i.e. mediate, postulates, because they require reflection on metaphysical externalities related to our pursuit of the moral law. Ostaric (2017): 105.
[313] *KdU* 5:469.
[314] *KdU* 5:467.
[315] *Willkürliche Erdichtung*, is Kant's German phrase. *KdrV* A822/B850.
[316] *KdrV* A820/B848. See Ostaric's distinction between persuasion and conviction in Ostaric (2017): 98.

second *Critique*, is a merely reasonable way of accounting for one's experience, which theoretical reason can neither prove nor disprove.[317]

The rational faith based on the moral law is very different from a mere credulity in a historical faith, i.e. the esoteric doctrines of a certain religion that are dogmatically asserted without any connection to reason.[318] Kant resisted Jacobi, who thought that Enlightenment rationalism leads to Spinozism, and so the best faith is one that transcends reason. Kant accuses Jacobi of allowing common reason to fall into an enthusiasm as it dethrones speculative reason, i.e. the intuitions of common sense are taken on faith in spite of a lack of evidence.[319]

Reason, to genuinely be reason, has to meet a publicity requirement that transcends embedment in particular cultural traditions, mere instrumentality in pursuit of private ends, and special modes of intuition that allows individuals to discern normative truths.[320] Reason can meet this publicity requirement in virtue of an organization around a single, central, and universally presupposed principle of the unconditioned.

The metaphysical agnosticism of epistemological TI means that reason may be frustrated in its positing of the unconditioned. Epistemological TI is unable to banish the specter of the Newtonian absolutization of space and time, as metaphysical TI does in virtue of the Antinomies, which show the contradictions that arise in virtue of a dogmatic

[317] For example, we might hypothesize a God to account for the order in the world. *KdpV* 5:142.
[318] *R* 6:103.
[319] Ibid., 8:134.
[320] I am referring to Onora O'Neil's account of reason, which by nature rejects arbitrary moves that might be associated with growing up in a certain culture, having certain subjective goals, or have special semi-mystical insights into moral facts. Onora O'Neil (2000): 11.

realism about space-time.[321] If we have not proven that Newtonian absolutization of space-time and its denial of substantial forms is not true, then it is possible that our practical cognition of ourselves is not a fact, but an illusion. The epistemic status of our commitment to absolute freedom would dwindle to a mere a priori opinion. Having undermined the practical cognition of absolute freedom, we could merely engage in a priori speculation about freedom that has no basis in experience or in analytically true definitions. In other terms, absolute freedom would dwindle to a persuasion, that is not binding on others: since Newtonian metaphysics may be true for epistemological TI, there is no demand for universal assent to absolute freedom.[322]

It is opportune at this point to engage in detail with Allison's account of the relationship between the intelligible and empirical characters. Epistemological TI accepts a metaphysical agnosticism that does not reject the possibility that space and time characterize reality in itself (this is the so-called neglected alternative[323]), and this prevents the guarantee of the possible existence of unconditioned absolute freedom as a genuine supersensible power. What is unconditioned cannot exist in the empirical domain of space and time: whatever is in time is *conditioned* by past *antecedentia*, and what is in space is conditioned by the different parts of space.[324] Allison nevertheless is intent on

[321] Karl Ameriks (2003): 107-111. The Antinomies teach us, according to Ameriks, not to treat our experience of spatiotemporality as an absolute measure.
[322] But, Kant proves the reality of absolute freedom from the presence of the moral law in the Analytic of Pure Practical Reason in *KdpV* 5:19.
[323] The neglected alternative stems from the view that establishing space and time as a priori conditions of knowledge does not guarantee that space and time do not also characterize reality in itself. See the discussion in Norman Kemp Smith (1962): 113-4. Guyer documents, and rejects, the interpretation of TI as a form of epistemological modesty that *cannot be sure* that things in themselves are as we represent them to be. Since epistemological modesty leaves open the possibility that things in themselves may be spatiotemporal, it does not reject the neglected alternative. Guyer (1987): 333.
[324] *KdrV* A412/B439.

articulating agency in terms of the so-called double character involving an intelligible and empirical aspect. The character is double in the sense that it involves both spontaneity at the intelligible level, and determinism at the empirical level. The metaphysics of transcendental idealism, I think, successfully resolves the tensions in integrating this double character, through the downgrading of the empirical character to a mere appearance. This downgrading means that the empirical character is not necessarily subject to the will of God, in the kind of theological determinism we find in the pre-Critical Kant. Nor is it the case that the appearance of the empirical character is ontologically identical to the intelligible character, and only different in terms of level of confusion, as we see in Leibniz. It is not the case that the agent is both determined and spontaneous in the identical ontological sense, but instead the agent is noumenally free in itself, but *appears* determined.

Allison resists the idea of the metaphysical commitment to a subject-relative empirical character, since he thinks that the idea that the intelligible aspect of reason appears as an empirical character permits an un-Critical inference from the empirical to the intelligible domain. Allison proposes the route of a "potentially attractive" form of compatibilism, which is predicated entirely on the empirical character. Kant mentions a "psychological concept"[325] of freedom, which dispenses with the idea of an absolute spontaneity incompatible with the empirical series. The psychological concept of freedom would involve a description of action in terms of maxims reconstructed on the basis of regular appearances of behavior associated with one's empirical character.[326] We

[325] Kant uses this expression in *KdrV* A448/B476.
[326] See the discussion in Allison (1990): 33.

are "free," in this compatibilist picture, insofar as we can describe our actions in terms of maxims, even though these descriptions are entirely naturalized reconstructions based on the empirical character.[327]

In spite of this possible route of attractive compatibilsim, Allison recognizes that the intelligible character is important to Kant as the proper arena for actions based on moral or pragmatic imperatives, i.e. that are not solely driven by non-rational impulses.[328] The spontaneity of reason allows it to prescribe its own order of ends independently of any empirical drives, and so it requires a space outside of the order of nature.[329] Allison's key move is to tie this spontaneity of reason not to a contra-causal power, but to spontaneity in an epistemic sense. The spontaneity of reason is reduced in Allison to transcendental apperception, which unites different representations into a single "I think" in a way that cannot be completely ascribed to a passive receptivity to external stimuli.[330] The empirical character subject to natural laws cannot fully account for the rational agent, because the agent *epistemically* has a unity of consciousness in virtue of which it self-consciously recognizes its reason as reasons. A rational agent cannot take a certain reason as properly participating in a categorical or hypothetical imperative on purely materialist-causal terms.[331] The spontaneous recognitional capacities of reason provide a merely

[327] Kant mentions how we can divine a rational rule animating an agent by way of regularity of effects of the empirical character in *KdrV* A549/B577.
[328] Allison attributes an incompatibilist position to Kant, against those who think Kant's moral psychology can be entirely based in a compatibilist framework. Allison (2008): 458.
[329] Kant mentions this spontaneous order of reason at *KdrV* A548/B576.
[330] Kant mentions the pure apperception at *KdrV* A547/B575.
[331] Allison discusses these issues in his paper, Allison (1989): 201.

formal and a priori framework, and so cannot themselves figure as empirical phenomena.[332]

By invoking an epistemic spontaneity, Allison is able to escape what he thinks is the "dogmatic metaphysical flavor" of a supersensible power.[333] The evaluative framework of "taking as" under the unity of apperception is a merely regulative ideal for rational agency, i.e. it prescribes formal relationships without ontological commitments.[334] Allison's version of freedom is merely practical freedom, which establishes an independence from necessitation by one's empirical impulses.[335] Allison's view of freedom does not involve, at least in terms of any ontological dependence, *transcendental* freedom, i.e. a spontaneous power of choice that lacks any causal *antecedentia* and which therefore cannot be a possible object of experience.[336] Practical freedom is only related to transcendental freedom in a conceptual manner, as transcendental freedom for Allison represents a kind of divine freedom independent from sensuous affection altogether, whereas practical freedom is mere independence from sensuous *necessitation*.[337]

The glaring problem, though, is that the merely epistemic spontaneity does not lift Kant out of the compatibilism of his pre-Critical phase, nor does it adequately separate

[332] Pure apperception is therefore distinguished from the temporal phenomena of inner sense. See *KdrV* B278.
[333] See Allison (1990): 39.
[334] See the distinction between regulative and constitutive principles at *KdrV* A179/B222.
[335] Kant is consistent in his definition of practical freedom in both the Dialectic and the Canon: *KdrV* A534/B562 and A802/B830.
[336] See Kant's definition of transcendental freedom at both *KdrV* A533/B561 and *KdrV* A803/B831. Allison notes that practical freedom is all that is needed for the Incorporation Thesis in Henry Allison, "Kant on Freedom: A Reply to My Critics," *Inquiry* 36 (1993): 443.
[337] Ibid., 444.

Kant from the determinism of the Leibnizian *spirituale*. The pre-Critical agent, though subjected to the will of God, may still take itself as acting on certain reasons. The pre-Critical agent can sustain a freedom that is not a metaphysical *explanans*, but merely integral to the *conception* of our agency.[338] Conceiving oneself as free lends itself to a pragmatic anti-realist account of freedom, in which we represent ourselves to ourselves *as if* we are free, to direct ourselves a certain way or to satisfy certain psychological needs, without an ontological commitment to the reality of freedom.[339] The Leibnizian agent, whose appearance of causal determinism characterizes its *an sich* nature as its confused appearance (i.e. it is not separately constituted by the form of human sensibility), may also pragmatically think of itself under a deliberative framework of self-posited ends. A merely formal version of an imperative, upheld by an agent that is epistemically spontaneous but ontologically heteronomous, is theoretically unstable, insofar as it represents a powerless illusion for an agent ultimately determined by her strongest desire.[340] The danger here is a kind of voluntarism (i.e. the will controls the intellect), in which the will is driven by empirical incentives in a way that is detached from a merely formal and powerless epistemic spontaneity of the intellect. Kant writes, "…If the freedom of our will were none other than…psychological and comparative but not also transcendental, i.e. absolute, then it would at bottom be nothing better than the freedom of a turnspit…"[341] What is needed to lift Kant out of this "turnspit"[342]

[338] Allison admits that his version of freedom has no explanatory power in Allison (1993): 461.
[339] Ostaric discusses this pragmatic view, and an anti-realist perspective based on psychological needs, in Ostaric (2017): 96-7. A pragmatic view is taken, for example, by a doctor unsure of a patient's illness, who nevertheless makes a pragmatic (i.e. non-ontological) judgment to guide her therapeutic interventions. *KdrV* A824/B852.
[340] This is the objection of Stephen Engstrom, which Allison discusses in Allison (1993): 446.
[341] *KdpV* 5:97.
[342] "Turnspit" freedom is a lesser kind of spontaneity Kant refers to as *spontaneitas secundum quid*. See Allison (1993): 445.

compatibilism is an intelligible character capable of genuine exercise of a metaphysical power. Allison's merely epistemically spontaneous agent still behaves heteronomously, insofar its ends are controlled by empirical incentives.

Inability to deduce the moral law as apodictically certain in epistemological TI. The unconditional moral law clashes with natural causes, since acting in response to a natural impulse can never match rational and impartial considerations of humanity as such. The incongruence between the unconditional law and natural causality can only be remedied by introducing another form of causality. A merely heteronomous agent, that only reacts to external stimuli, is not capable of acting according to moral law that maintains its standards regardless of empirical incentives. In epistemological TI, we can posit the absolute freedom required by the moral law only as a non-ontological standpoint, given that it might be illusory in the light of the absoluteness of theoretical cognition.

If we merely naturalize the will and detach it from the demand of an unconditional moral law that pertains to pure rationality as such, it is not clear how we are to establish a principle of this naturalized will as normatively privileged. A pure principle of reason as such (i.e. it is not a mere quirk of a single individual's reason) has a moral authority that a mere drive, based on empirical stimuli or imposed by an individual reason as a private principle, does not. Epistemological TI would undermine the status of the unconditional moral law, by permitting the possibility that the human person has no pure reason, but is only subject heteronomously to empirical inclinations.[343] An agent

[343] See Karl Amerik's critique of Karl Reinhold's degradation of Kant's idea of a pure will in touch with an unconditional moral law into a merely naturalistic drive. Karl Ameriks (2012): 194.

capable of "being a law to itself,"[344] i.e. following a principle of reason as such apart from any empirical incentives, must have a motivational independence in virtue of which it can set its own ends, and not merely find means for ends pre-determined by empirical motivations.[345] A heteronomous agent can only respond conditionally according to the ends circumstances dictate for her, and so cannot be subject to unconditional duty.[346]

Conclusion

In this chapter, I explained how a diversification in grounding relationships allow Kant to escape the compatibilism of his pre-Critical period. By establishing the temporal series as metaphysically conditioned, the human agent is no longer subject in its very actions to the will of God. Instead, actions exist in the empirical series, which God does not create, but which are only secondary effects having to do with the way the person appears. I also devoted a section to differentiating the Kantian appearance/thing in itself distinction from the Leibnizian one, showing how the form of space and time for Kant have a role in constituting a distinct appearance, and are not merely degraded and confused versions of things in themselves. I then showed how epistemological TI is unable to secure the real possibility of the unconditioned, and this has a host of negative consequences, including downgrading our absolute freedom from a cognized fact to a

[344] *GMS* 4:440.
[345] Allison acknowledges the motivational independence of transcendental freedom, and its importance in deducing the moral law, in Henry Allison (1991): 11.
[346] Kant makes this argument in *MDS* 6:385. Derk Pereboom is a notable philosopher who argues—mistakenly, I think—that even though science disproves moral responsibility, we can still have morality, meaning, and value. Pereboom (2001): xiv.

mere subjective persuasion, and also imperiling the applicability of an unconditional moral law to us, who may in the end be merely heteronomous creatures.

The idea that free will acts as a kind of substantial form for empirical events, structuring the empirical laws that govern them, may offend those who insist on a strong noumenal ignorance for Kant, insofar as my view definitely connects a set of empirical actions to a single individual. It appears, then, that I am invoking a one-object view of the appearance/thing in itself distinction. One might object that this one-to-one correspondence shows too much knowledge of the uncognizable noumenal realm. One might also claim that a one-object view cannot bridge the gap between spatiotemporality and non-spatiotemporality. What is spatial and temporal cannot be one object with what is neither.[347] As for the first epistemic claim, we can establish one-to-one correspondence, at least with respect to empirical actions, in virtue of the practical, but not theoretical, cognition of the agent as absolutely free. The agent's cognition of itself as practically free allows it to establish itself as the unconditioned source of its empirical actions. As for the metaphysical claim, of one person being both spatiotemporal and non-spatiotemporal, there is nothing incoherent about this view. Spiritual people may think, for instance, that God sees all our actions at once, in light of His foreknowledge, and so, in terms of our intelligible character, we are not in time. Moreover, we may have a soul that is immaterial and does not occupy space.

Those who promote epistemological TI may object that they can secure a place for the unconditioned, in terms of an epistemic distinction between modes of

[347] Nicholas Stang makes this objection in Nicholas Stang (2014): 111.

consideration of an object. There is a non-spatiotemporal aspect of the object in virtue of an abstraction from our sensible conditions.[348] However, the question of what really exists, independently of perspectives we might identify, persists. Epistemological TI never commits itself to the idea that the spatiotemporal world is not an absolute totality, since we can take a non-spatiotemporal perspective with our pure understanding even if the spatiotemporal world is absolute. Epistemological TI therefore leaves open the possibility that our freedom is causally inert, merely a way of talking, that betrays the early metaphysical commitments of Kant to a form of potential interaction between substances that precedes any spatiotemporal events.[349] This is not to say that epistemic conditions are not important for freedom. They are, since free choice cannot occur if one is not conscious of oneself as so choosing. But, a metaphysical basis for free choice is also very important.[350] Recasting freedom as a mere perspective represents the undermining of our ability to present the postulates of practical reason, following from the moral law, as *theoretical* propositions that have ontological reference—i.e. they are about the world—even though they are theoretically indemonstrable.[351]

[348] In Hud Hudson's theory, a teleological and deliberative aspect of agency is attained by abstracting from the laws that characterize the physical domain, but not the mental domain, in spite of a token-token identity between the mental and the physical. Hud Hudson, (1994): 66.
[349] See *Diss* 2:390, where Kant says that the form of potential causal interaction has to be what unites the world, and not actual causal interaction.
[350] Fischer identifies both epistemic and "freedom-relevant" conditions of moral responsibility, cited in Mele (2006): 76.
[351] See Guyer's discussion of practical postulates in Guyer (1979): 145.

CHAPTER 3

THE IMPLICATIONS OF THE FIRST AND THIRD ANTINOMIES FOR THE NATURE OF TRANSCENDENTAL IDEALISM AND, BY EXTENSION, THE METAPHYSICS OF AGENCY

In the Antinomies, Kant identifies contradictions into which reason falls as it tries to consider cosmological questions having to do with the magnitude and duration of the world, whether or not there are simple (i.e. partless) components of the world, whether or not there is causality according to freedom, and whether or not there is a necessary being.[352] The Antinomies deal in particular with reason's attempt to integrate its aspiration towards the unconditioned in the sequence of appearances—e.g. in seeking a limit to the spatiotemporal extent of appearances, or identifying a first cause of a causal series across space-time.[353] The Antinomies represent a crisis of reason, because they are wholly natural to it—it is not optional for reason to fall into self-contradiction, but instead the crisis occurs in light of sound arguments.[354] Kant identifies two dire potential outcomes of the Antinomies, which he seeks to avoid: a skeptical hopelessness, in which reason gives up on its aspiration towards the unconditioned and surrenders to a blind fideism, or a dogmatic stubbornness, in which reason rigidly adopts one side of the antinomial conflict without considering the equally meritorious opposing position. The

[352] The Antinomies begin at A405/B432.
[353] In this way, the Antinomies are distinct from the Paralogisms, which deal with the seat of representations in the soul, and the Transcendental Ideal, which has to do with the conditions of possibility in general. *KdrV* A406/B433.
[354] *KdrV* A421/B449 and A407/B434.

crisis of the Antinomies is driven by the stance of transcendental realism, and is supposed to be resolved by the new paradigm of transcendental idealism.[355]

The resolution of the antinomial conflict, therefore, is clouded by the controversy, which I discussed in chapter one, having to do with the different interpretations of both transcendental idealism and transcendental realism. The controversy hinges on whether the TR/TI distinction is one involving two differing epistemic standpoints, i.e. a God's-eye ability to gain cognition of reality in itself and a modesty restricted to the conditions of human sensibility, or whether the distinction involves two differing metaphysical positions, i.e. whether the spatiotemporal series of empirical cognition is supposed to be taken as metaphysically absolute or as metaphysically conditioned.[356] In short, the controversy over how to resolve the antinomial conflict is between an approach that emphasizes methodological restrictions on what we can know—i.e. to accept as undecidable cosmological questions that go beyond the limits of what can possibly be a sensible intuition[357]—or an approach that involves a new metaphysical claim, namely, that reason's aspiration for the unconditioned must take place in a metaphysically fundamental domain of things in themselves that are not in space and time.[358]

The antinomies therefore represent an important test for the nature of TI. This test of the nature of TI will provide illumination for the kind of framework TI might provide for agency. The third antinomy represents reason's attempt to posit the unconditioned in

[355] *KdrV* A491/B519.
[356] Henry Allison identifies these differing approaches in Allison (2004): 394.
[357] Paul Guyer identifies this methodological solution to the Antinomies in the early works of the 1770's such as the *Inaugural Dissertation*. Paul Guyer (1987): 387.
[358] Paul Guyer thinks that this so-called metaphysical dogma is the real solution of the Antinomies, much to his disappointment and that of other commentators who seek to find a methodological resolution. Ibid., 385.

the form of transcendental freedom, i.e. an absolute causal spontaneity able to initiate a new sequence of events.[359] In the first section of this chapter, then, I will discuss different interpretations of the TR/TI distinction that is supposed to be the key presupposition both driving and resolving the antinomial conflict. In the second section of this chapter, I will test the ability of epistemological TI, which emphasizes cognitive limitations without introducing a shift towards metaphysical idealism of the spatiotemporal world, to resolve the antinomies. In section three, I will discuss how the metaphysical contingency of space-time, which I think is essential to resolving the antinomial conflict, is a view that has roots in Kant's early scientific work as well as other pre-Critical philosophical work. In section four, I will discuss how a failure to establish space-time as metaphysically non-fundamental leads to a weak perspectival view of agency for epistemological TI. Throughout, I will address only the first and third antinomies because these bear most directly on the metaphysics of agency that is the central project of this dissertation. I will argue that the upshot of the first antinomy is that we cannot define an absolute totality—i.e. a series of conditioned-condition relationships leading to something unconditioned[360]—in terms of space and time. An attempt to posit the world, the sum total of appearances in space and time, as a metaphysical totality, leads to the contradiction of a world that is both infinite and finite. We work ourselves out of this contradiction by limiting space and time to the a priori structure of the human subject, so that we can posit the unconditioned outside of the spatiotemporal framework. The restriction of the spatiotemporal framework to the empirical synthesis of the human subject, and the corollary of a lack of complete magnitude, carries over to the third

[359] The third antinomy begins at *KdrV* A444/B472.
[360] Kant defines a totality at *KdrV* A409/B436.

antinomy in the withdrawal of the causal structure of the empirical series from the unconditioned reality of the human agent. The fact that the causal structure in space and time does not characterize the agent in her unconditioned reality permits the possibility of a power of freedom that is undetermined by empirical factors. My interpretation of the first antinomy, then, as well as its application to the third antinomy, does not consider the first antinomy as a mere dispute about the relationship between the world and space-time—i.e. whether the world is a finite entity embedded in space-time or is an infinite entity coextensive with it. Instead, my interpretation shows the implications of the first antinomy for the nature of space and time itself, namely, that they are not subject to absolutization as a finite or infinite totality.[361] It is the very attempt to articulate the unconditioned in terms of the temporal series, which is a series of conditioned things arising from antecedent grounds, and the spatial series, which is an aggregate of different parts mutually conditioning one another as boundaries[362], that leads to the contradiction of the finite-infinite world.

Different Interpretations of the TR/TI Opposition

Henry Allison is famous for introducing an interpretation of TI and TR as two metaphilosophical, or meta-epistemological, perspectives, rather than two ontological doctrines. TR is a standpoint in which we take a God's-eye perspective on the totalities

[361] Allison claims that the first antinomy is only about the relationship of the world to space-time, not the nature of space-time itself. Allison is committed to avoiding any ontology of space-time, since the epistemic thrust of TI for Allison emphasizes only the role of space-time in cognition. Allison (2004): 366.
[362] For a discussion of the conditioned realities in space and time, see *KdrV* A412/B439.

involved in the cosmological questions. This God's-eye perspective disregards the spatiotemporal structure of human sensibility, which gains cognition of things only as localized in particular parts of the spatiotemporal fabric. TI, on the other hand, is an approach of epistemological modesty that remains centered in the localized spatiotemporality in which sensibility can cognize things, and so remains non-committal and agnostic about larger cosmological questions having to do with the world as a totality.[363] The shift from TR to TI, in this interpretation, involves "a bit of metaphilosophical therapy" that shifts us from grand metaphysical claims that abstract from our cognitive limitations, to more modest and localized claims within human epistemic limits.

It must be noted that this metaphilosophical shift has nothing to do with metaphysics. Allison explicitly distinguishes the metaphilosophical shift of TI from a metaphysical shift that adopts a new phenomenalism about empirical objects. Staying within an anthropological perspective, that respects our cognitive limitations to localization in space-time, does not have any entailments for the issue of a metaphysical idealism about space-time, which ties its existence exclusively to the subject and denies its characterization of reality in itself.[364]

Carl Posy is another commentator who views the shift from TR to TI in the antinomies as a methodological shift. Posy differentiates between TI and TR in terms of a theory of meaning: TI restricts meaning to that which we have evidence to assert, whereas TR ties meaning to states of affairs in the world that are independent of our

[363] Henry Allison (2004): 394-5.
[364] Kant discusses a formal idealism about space and time, which permits a realism about the objects in space and time, in a passage cited in Paul Guyer (1987): 414.

evidential capacities. We adopt TI, to resolve, for instance, the contradiction about the magnitude and duration of the world in the First Antinomy, to establish that we cannot make meaningful claims about the world as a totality, because we could never have evidence of the boundaries—or lack thereof—of the entire world. It is TR, that tries to establish some truth about the world independently of the evidence available to us, that leads to contradictions in the antinomies.[365] Similarly, for Allison, the God's eye view seeks to establish metaphysical truths that transcend the evidence available to human limitations, whereas the metaphilosophical therapy of TI limits us to claims based on evidence we can obtain.

Ameriks, on the other hand, thinks that the Antinomies are meant to show the problems of a dogmatic realism about space and time. Dogmatic realism is actually the problem of the so-called neglected alternative, which is the view that claiming that space and time are the a priori conditions of sensible intuition does not guarantee that space and time do not also characterize reality as such, i.e. things in themselves and not merely as they appear.[366]

If we are dogmatic realists about space and time, we think they track reality in itself, in its absolute dimensions detached from any human perspective. The presupposition of dogmatic realism about space and time means that, when reason thinks the unconditioned, it has to locate it in space and time. So, in considering the cosmological question of the magnitude and duration of the world, the dogmatic realist rules out an infinite world because this would require an infinite synthesis requiring an

[365] Carl Posy (1983): 82-84.
[366] Allison discusses the neglected alternative objection in Henry Allison (2013): 314.

infinite amount of time.[367] At the same time, the world cannot be finite, because this would require an arising in *empty time,* which is undifferentiated and so provides no basis for an arising at a certain point in itself.[368] Since thinking the unconditioned in space-time leads to this contradiction, the proper resolution is to deny the dogmatic realism about space-time. Once we accept an idealism about space-time, reason does not have to think the unconditioned in it, and the contradictions disappear.[369] The proper way to resolve the antinomies, then, is a metaphysical shift that relegates space-time to the status of a mere appearance associated with the human subject, so that reason can think the unconditioned in the non-spatiotemporal domain of things in themselves.

What makes Ameriks's resolution of the antinomies different from that of Allison is that, in the former, there is a definite metaphysical position taken on the nature of space and time. Allison proposes only a "meta-philosophical" shift, in which we check the transcendental illusion by which we posit the totality of the conditioned-condition series as something really given. The transcendental illusion, as defined by Michelle Grier, Allison's student, emerges in rationalistic metaphysics in the form of a hypostatization of reason's demand for complete explanation, so that this demand becomes an ontological commitment to the real givenness of an unconditioned totality.[370] We can walk back this transcendental illusion, and cease to posit the world as a totality, and so seem to win a victory against the contradictions of the rationalist metaphysics underlying the positions in the antinomial conflict. However, merely walking back from the transcendental

[367] *KdrV* A429/B457.
[368] *KdrV* A427/B455.
[369] See the discussion in Karl Ameriks (2003): 107-111.
[370] Michele Grier (2001): 2.

illusion in a meta-philosophical turn against rationalist metaphysics does not eliminate the specter of the neglected alternative. Space and time might continue to characterize objective reality, even as we refrain from considering the world as something absolutely given. The objective reality of space-time would prohibit the possible existence of the unconditioned.

Allison himself insists that things in themselves are not spatiotemporal. We can distinguish two senses of things in themselves: an ontological sense, in which they are different objects from sensible objects, i.e. noumenal non-sensible objects, and an epistemological sense, in which they represent a detachment from epistemic conditions associated with human knowers.[371] Allison says that non-sensible noumenal objects cannot be spatiotemporal, because otherwise they would be possible objects of experience, i.e. sensible. Also, things in themselves in the epistemological sense cannot be spatiotemporal, because they are non-spatiotemporal by definition—we are considering them in detachment from the conditions of human cognition.

However, without the metaphysical shift of Ameriks, in which we limit the spatiotemporal framework to the human subject, the neglected alternative of the metaphysical ultimacy of space and time is a live possibility. We can walk back our transcendental illusion, and cease to regard the causal series over time in the third antinomy as a given totality. Absolute freedom, we might also argue, is a non-sensible object that is by definition not in space and time. The problem, though, is that definitional stipulations, of what is detached from human epistemic conditions and so tautologically

[371] Henry Allison (2013): 317.

non-spatiotemporal, fails to secure the possibility of the existence of a non-sensible absolute freedom as a real object. Space and time may characterize the fabric of reality as something ontologically fundamental, even if we make the trivial claim that what is independent of human conditions of cognition (i.e. space and time) is not in space and time. Ameriks's resolution involves the recognition that spatiotemporal appearances, for Kant, are necessarily conditioned, insofar as they mediated by our cognitive structures and not simply present to us as brute facts. If spatiotemporal appearances are necessarily conditioned, then reason's search for the unconditioned must take us beyond the empirical domain of appearances.[372] Allison is not able to guarantee the possibility that this non-empirical unconditioned thing might exist. A mere recognition that we are cognitively limited to what is in space-time does not guarantee that any search for the unconditioned will be protected from an ontologically fundamental spatiotemporal domain that is pervasively conditioned.

Ameriks moves from the Transcendental Aesthetic, which establishes space and time as a priori intuitions of sensibility, to a completion in the antinomies of the Transcendental Dialectic, because the latter finally completes the dismissal of the neglected alternative. Space and time cannot be ontologically fundamental because reason cannot think the unconditioned in space and time without contradiction.[373] If we remain within the position of the Transcendental Aesthetic, i.e. that space and time are a priori intuitions of human sensibility, we cannot banish the neglected alternative, but can only make the tautological claim that what we cannot cognize is necessarily not in space

[372] Karl Ameriks (2006): 149-150.
[373] This is the long argument for TI that Ameriks mentions in Ameriks (2003): 136.

and time. This limitation of cognition does not mean that all of reality is not spatiotemporal, and a pervasively spatiotemporal reality suppresses the unconditioned.

Does Epistemological TI, With Its Emphasis on Cognitive Limitations, Really Resolve the Antinomial Conflicts?

In this section, I will show that the metaphilosophical resolution of commentators like Allison fails to resolve the antinomial conflict. At best, epistemological TI provides only the unsatisfactory resolution of being able to posit the unconditioned as a non-committal standpoint that does not track reality, and fails to secure the unconditioned as something possibly really given. First, I will articulate how the metaphilosophical approach of Allison is supposed to resolve the conflict. Then, I will introduce four objections to the metaphilosophical/therapeutic resolution: 1) the TR-P_2 combination Allison claims is at the heart of the antinomial conflict actually does not by itself drive the conflict, and therefore its removal does not fully resolve the antinomial conflict, 2) an empirical whole in the spatiotemporal domain is not intrinsically impossible, 3) the metaphilosophical resolution fails to protect reason's aspiration towards the unconditioned, and 4) the metaphilosophical resolution fails to adequately account for the difference in resolution between the mathematical and dynamical antinomies.

The resolution of Allison involving the rejection of the TR-P_2 combination: Allison's diagnosis of the antinomial conflict is that it involves a commitment to a realism about the metaphysical givenness of the cosmological series of the world. The

TR-P_2 combination is a combination between transcendental realism, and the premise that, if a conditioned thing is given, the whole series of conditions leading up to it, along with its unconditioned origin, is also given.[374] P_2 is the premise of the transcendental illusion: it involves a hypostatization of the merely logical and subjective principle of reason to find the unconditioned for the conditioned cognitions of the understanding.[375] TR is a metaphilosophical perspective in which we take a God's-eye view of reality, in detachment from the conditions of human sensibility, and make metaphysical claims not merely about objects of our experience, but about objects in general.[376]

We get the antinomial conflict when the toxic TR-P_2 combination leads us to posit the world as a totality—i.e. we posit the entire series of conditions leading up to the unconditioned from a God's-eye perspective, taking a stance on the nature of the world in itself, apart from the limitations of our human perspective. So, in the first antinomy, we take a stance on the totality of the world in terms of its entire magnitude and duration, and end up with the contradiction of a world that is both finite and infinite. In the third antinomy, we take a stance on the totality of the causal series in the world, claiming that there is both a spontaneous first cause and that the causal series is an infinite series of conditioned-condition relationships.

The contradictions are unsustainable, and the recommendation of Allison is that we take a stance of epistemological modesty. Instead of the toxic TR-P_2 combination, TI is form of therapy that forces us to stay within the limits of our human sensibility, rather

[374] See the discussion at *KdrV* A307-8/B364.
[375] Henry Allison (2004): 329. Allison states that the transcendental illusion takes the mere thought of the unconditioned/totality, and infers that it is a whole cognizable by a human being. Allison (2001): 321.
[376] See Allison's discussion of the theocentric perspective at Allison (2004): 27.

than venturing with our pure understanding to make grand metaphysical claims about the totality of the world in itself. With the adoption of TI, we restrict ourselves to a piecemeal empirical synthesis at precise locations in space-time, and within this anthropocentric perspective, we cease to take any metaphysical stances on the questions of cosmological totality. We no longer have to worry about contradictions within this limited anthropocentric domain.

There is certainly truth in claim that Kant does not want us to make grand metaphysical claims independently of possible experience. Part of the Critical philosophy is a methodological restriction of the pure concepts of the understanding to experience, and this would rule out any totalizing tendencies that seek to expand this experience into the unconditioned limits of the nature of the world in itself. The idea that we are not to mistake the subjective conditions of sensible cognition for objective principles of the existence of things in themselves goes back to the subreptic axioms of the *Inaugural Dissertation*.[377] Subreptic axioms represent Kant's opposition to the idea that we are to project what we know are conditions for sensible representation, in a totalizing way, on to the ontological nature of reality in itself. Subreptic axioms naturally lead to the idea of the noumenon[378], which is a sort of cognitive black hole meant to set a limit to the tendency of sensibility to project itself on to the ontological structure of things in themselves.

Objection no. 1: the TR-P_2 combination does not by itself entail the antinomial contradiction. The first problem I note with this purported resolution is that the TR-P_2

[377] *Diss* 2:412. For an excellent account of subreptic axioms, see Michelle Grier (2001): 57.
[378] *KdrV* A255/B311.

combination does not itself entail any contradiction, between finitude and infinitude, or a first cause or an infinite conditioned-condition series. One can embrace the TR-P_2 combination, and claim to have insight into the nature of the world as a totality from a God's-eye perspective, and not run into any contradictions. One could, from perspective of the TR-P_2 combination, make the dogmatic claim that the world is finite, or that there is a first cause. There is no entailment, from the mere taking of a God's-eye perspective, that one must end up in contradictions about cosmological questions. For example, we could imagine a cosmological debate about the magnitude of the world that did not degenerate into contradiction. The rationalist position, let us say, that the spatiotemporal world is finite, is a valid and necessary argument, while the empiricist position of infinitude is invalid. There would be no contradiction here, but the rationalist position would still represent a TR-P_2 combination, insofar as it posits the totality of the world as really given. We would be urged to think the totality of the world not as something really given, but in a non-propositional, or non-assertoric, way, i.e. as a mere mode of consideration. The real givenness of a metaphysical totality is not as such the problem of the antinomial conflict, therefore, since one can hypothesize a metaphysical totality that does not involve a contradiction. The contradiction, instead, of the antinomes has to do with the frustration of reason in positing the unconditioned *in space and time*, which requires a recognition, as a solution, that, metaphysically, space and time are conditioned realities.

Accordingly, if the TR-P_2 combination as such does not drive the antinomial conflict, merely walking back from it, by taking a stance of epistemological modesty, will not resolve the antinomial conflict. Reason can continue to make sound arguments for

contradictory positions in the antinomial conflict, and suffer a self-destructive entanglement, even if we avoid cosmological questions as uncognizable given the limits of our sensibility. The issue here has to do with the kind of faculty in which the antinomial conflict is rooted.[379] Cosmological questions certainly transcend our ability to have cognition, which, for Kant's discursive intellect, requires a union of the concepts of the understanding and intuitions of the sensibility.[380] Nevertheless, we can still think about cosmological questions with our pure reason. Thinking, unlike cognition, only requires a lack of contradiction, and does not require an object to be given in sensibility.[381] The recommendation of epistemological TI is that we should refrain from making assertoric claims about the real givenness of the world, as something either finite or infinite, and instead retain the thesis and antithesis positions as perspectives without ontological commitments that we can take up in different contexts. But, merely moderating the claims of the thesis and antithesis arguments into non-propositional stances that do not have to do with the nature of things[382], does not rescue reason from a permanent frustration in its attempt to posit the unconditioned as a possibly existing object. Moderating the transcendental illusion does not preclude the possibility that reality in itself is spatiotemporal, and that the unconditioned is permanently supressed, as

[379] Allison notes, with respect to the antinomy of taste, that a genuine antinomy cannot arise in relation to what is merely agreeable. There is no contradiction in the fact that what seems agreeable to one person is not agreeable to another, because the claim to agreeableness has mere private validity and is not a claim to universality. Reason creates an antinomy in the first *Critique* because it makes a claim to universality. See Allison (2001): 237.

[380] Kant's famous doctrine in the Transcendental Logic is that intuitions without concepts are blind, and concepts without intuitions are empty. *KdrV* A51/B75.

[381] See the distinction between cognition and thinking in the B-Preface, at *KdrV* Bxxvi.

[382] One can see this non-propositionality in full display in Allison's version of free will, in which he claims that freedom is a mere perspective that is not a "fact of the matter." Allison (2004): 49.

something ontological (and not merely perspectival) by prior conditions in time and boundedness in space.

The metaphysical approach to the resolution of the antinomies does not suffer from the problem of the epistemological approach, namely, that what it identifies to be the key premise essential to both driving and resolving the antinomies, i.e. the transcendental illusion, is really logically independent of any contradiction. The premise crucial for the antinomial conflict in the metaphysical interpretation is not a merely overly ambitious claim to cognition, but a failure to debunk the neglected alternative. A contradiction occurs, that is, in reason's search for the unconditioned, not because of an assertoric claim of real givenness of a totality as such, but because this search is forced into a spatiotemporal framework. The ascription to the spatiotemporal framework of an unconditioned totality leads to a contradiction, and so the spatiotemporal framework cannot be the proper arena of the unconditioned—it must instead only characterize a conditioned, subjective domain of the receptive intellect.

The TR-P_2 combination is not committed to a dogmatism about space and time. Dogmatism about space and time means that reason's aspiration towards the unconditioned must unfold in an ultimate reality characterized by space and time.[383] The TR-P_2 combination, in fact, merely establishes an absolute totality of conditions bottoming out in something unconditioned, in a way that transcends human cognitive

[383] Karl Ameriks notes that the unconditioned cannot unfold in the empirical framework of space and time. What is given in experience is conditioned for Kant—it does not simply exist in the manner of a Humean impression. The synthetic activity of our minds in receiving impressions, through our a priori structures, by which appearances are conditioned, means that there is an unconditioned element of what appears to us that is non-empirical. We abstract from the aspect of an appearance conditioned by our minds to reach its unconditioned and non-empirical ground. Karl Ameriks (2006): 150. Ostaric makes the same point, i.e. that appearances are *given* to the mind, and not merely present, in Ostaric (2009): 162.

limitations. The establishment of an absolute totality can even involve an idealism about space and time. Allison characterizes Berkeleyan idealism, in which space and time are merely phenomenalistic objects, as a form of transcendental realism.[384] In other words, one can adopt the TR-P_2 combination, even as a Berkeleyan idealist that takes spatiotemporal objects to be objects of inner sense, in virtue of the fact that these psychological objects are taken as a metaphysical totality independent of the conditions of possible experience.[385] Allison also identifies the non-fundamentality of space-time in ontological interpretations of TI[386] as a form of transcendental realism, since it posits at the foundation of experience an absolute domain of things in themselves. This absolute domain is outside of us in a transcendental sense, and so metaphysical TI is guilty of a sort of God's-eye establishment of a totality that Allison thinks is the opposite of genuine TI.[387]

If, however, the TR-P_2 combination drives *any* metaphysics of totality, it cannot drive the contradiction of the antinomial conflict. Avoiding commitment to a metaphysical totality *as such* would mean abandoning metaphysics even regardless of the appearance of contradiction. Even a single uncontradicted thesis of an idealism about space and time, in which the world, as a mental object, is neither finite nor infinite, would

[384] Allison (2004): 26.
[385] Henry Allison differentiates between empirical and transcendental interiority (*in uns* and *ausser uns*) in Allison (2004): 24. The discussion in Kant takes place in *KdrV* A373.
[386] It needs to be noted that the non-fundamentality of space-time in metaphysical TI is importantly different from a Berkeleyan idealism about space-time. It may be the case that Berkeley does not posit space-time as an absolute entity existing independently of human sensibility. Kant in the *Prolegomena* says he is in agreement with Berkeley in denying that space and time characterize things in themselves—like Berkeley, he thinks they characterize only appearances. However, Berkeley is an empiricist about space and time, which means we acquire our knowledge of them in a piecemeal way through experience, whereas Kant thinks that space and time are a priori transcendental structures. *P* 4: 374-5.
[387] Allison (2004): 47. Guyer makes the similar claim that the supersensible substratum in Kant's theory of taste is an example of transcendental illusion. Paul Guyer (1979): 7.

count as an illicit TR-P_2 combination. Since TR-P_2 does not by itself generate a contradiction, merely abandoning metaphysics by abandoning TR-P_2 does not mean that reason ceases to generate valid and necessary, but contradictory arguments, when it tries to posit the unconditioned as a real object in space-time. If we cannot make assertoric claims about the unconditioned, as actually pertaining to a reality and not a mere perspective, we cannot be sure of the possible existence of non-sensible objects, such as God or absolute freedom. Reason confronts a spatiotemporal framework that forces its unconditioned into contradiction and non-existence, and so space-time cannot partake of the unconditioned—it is subject-relative. Epistemological TI cannot debunk the possibility that all of reality is spatiotemporal, and so cannot secure the unconditioned as something that can possibly exist.

When we walk back the dogmatism about space and time, we are able to get a genuine resolution that allows reason to have satisfaction of its need for the unconditioned in a really given object. If the spatiotemporal series is a mere appearance, and does not characterize reality in itself, then reason does not have to posit the unconditioned limits of the world in space and time. The series of appearances, therefore, is neither finite nor infinite.[388] Reason also does not have to posit the unconditioned first cause in the time-series, because, again, it is a mere appearance and not an absolute reality, and so we do not have to paradoxically establish the unconditioned first cause as itself an event in time that requires a prior event to account for it.[389]

[388] Kant describes this resolution at *KdrV* A505/B533.
[389] See the resolution of the third antinomy at *KdrV* A506/B534.

The solution of the Antinomies in the metaphysical approach, involving a non-ascription of unconditioned totalities to the spatiotemporal domain, is not dogmatic or skeptical. It is not dogmatic, because this would require a metaphysical insight into a cosmological question that would violate Kant's prohibition on the transcendental employment of the categories for cognition. A dogmatic solution takes a definite side, claiming, for instance, that the world is definitely finite, or definitely infinite, in virtue of some insight into fundamental ontology. The solution of metaphysical TI is not skeptical either[390], since a skeptical solution claims a dogmatic insight into fundamental ontology that validates two contradictory theses as, paradoxically, both true. The critical solution, which Kant adopts, does not claim dogmatic insight into either the thesis or the antithesis, since it validates neither of them, and so can neither take a side nor establish a contradiction.

The critical solution of metaphysical TI exposes the premise that reason must think the unconditioned totality of the world in the spatiotemporal and categorical structure of possible experience. The exposure of this premise reveals that the world, as the sum total of appearances, is *neither* finite or infinite, because its status as a mere appearance means it is not amenable to an unconditioned metaphysical status as a totality. Appearances are not determinate, insofar as they cannot exist as an absolute totality, because they are only given in the piecemeal progression of the empirical synthesis.[391] In the third antinomy, metaphysical TI also does not take a dogmatic stance, either in the definite affirmation of absolute freedom or in universal mechanism. Instead,

[390] Kant provides a list of three possible solutions to the Antinomies at *KdrV* A388, and later at *KdrV* B434.
[391] *KdrV* A499/B527.

metaphysical TI takes the critical solution, by showing how absolute freedom and mechanical causality may both be true (not the dogmatic affirmation both are true).[392] The metaphysical idealism about space-time is essential to showing how a universal mechanism of the causal order of appearances does not preclude absolute unconditioned freedom, because appearances cannot track the absolute nature of things and must only exist in relation to the empirical synthesis of the human subject.

Epistemological TI refrains from the claim of dogmatic insight into cosmological issues we see in both the dogmatic and skeptical solutions. However, it does not fully embrace the critical solution, either. Epistemological TI does not expose the premise that reason has to posit the unconditioned within space and time, because TR, which drives the antinomies, involves a dogmatic realism about the structure of space and time. Epistemological TI does not attack the unconditioned or conditioned metaphysical status of space-time, but instead the epistemic overreach represented in cosmological speculation. Epistemological TI does not tell us that the empirical series in possible experience is indeterminate, ontologically dependent on the empirical synthesis by the human subject, and so not part of the ontological structure of things in themselves. Kant writes that antinomies are resolved through the clearly ontological claim that appearances, "if they constitute a series, *exist* in the successive regress but otherwise do not *exist* at all."[393] Merely establishing space and time as epistemic conditions and restricting cognition to a limited anthropocentric sphere cannot have any bearing on to what extent appearances *exist* independently of the sphere of cognition.

---

[392] The difference between mathematical and dynamical antinomies is discussed at KdrV *A534/B562.* See Kant's conclusion that absolute freedom is logically possible at *KdrV* A558/B586.
[393] *KdrV* A506/B534, italics mine.

Epistemological TI resolves the contradictions of the antinomies, but only in virtue of a modification of the claims involved in them, involving a shift from assertoric claims to non-assertoric claims, i.e. claims that are not about reality but merely reflect certain standpoints. Absolute freedom, reconceptualized as a standpoint and not an assertoric fact of the matter, does not contradict universal subjection to the causal law qua standpoint. Allison claims that the order of ends subject to the moral law, in which practical reason takes itself as freely self-determining, is a regulative ideal akin to the theoretical analogue of the systematic unity of empirical cognitions.[394] The regulative status means that freedom is has to do with a formal framework within which we adjudicate our decisions in the light of the moral law, which does not have any constitutive role in terms of ontological content.[395] Two regulative ideals, of both absolute freedom and determinism, do not contradict one another, since they do not involve ontological commitments about the nature of the world. Though the contradiction of the antinomial conflict is resolved in this manner, it comes at the expense of being able to posit absolute freedom as a real object (i.e. constitutively, not regulatively). If we were to make the assertoric claim about our absolute freedom, we could not be assured that it would not be subject to, and eliminated by, the causal conditions of experience, which epistemological TI fails to establish as ontologically non-absolute.

Given the fact that the neglected alternative about an absolute space and time that characterizes reality in itself is never definitively removed by epistemological TI, practical reason is forced to contort itself into a merely regulative and non-assertoric

[394] Allison (1990): 40.
[395] Kant distinguishes between regulative and constitutive principles at *KdrV* A179/B222, and A509/B537.

faculty in order to avoid contradiction. Kant is clear, though, that practical reason is capable of assertoric cognition just like theoretical cognition. The concept of absolute freedom is in fact the "keystone of the whole structure of a system of pure reason, even of speculative reason," insofar as freedom, the existence of which is proven in virtue of the moral law, provides a firm epistemic foundation for the objective reality of the postulates of God and immortality.[396]

The mere recognition of space and time as epistemic conditions, and the modesty in restricting cognition to what is spatiotemporally localized, does not prevent the paradoxical existence of a world that is both finite and infinite taken constitutively, or an unconditioned causality that is nevertheless subject to a causal law connecting it to a prior event. Allison thinks that Kant abjures other views of space and time as establishing a metaphysics of space and time. The Critical approach, according to Allison, is to avoid all metaphysics of space and time, and establish them strictly as epistemic conditions (i.e. alternatives to ontology).[397] But, an exclusive focus on epistemic conditions amounts to metaphysical neutrality. The spatiotemporal series could exist as an absolute structure of reality, if we merely establish we can only have cognition of what is in space and time. So, establishing space and time as epistemic conditions does nothing to avoid a situation in which reason posits its unconditioned, as a real object, within the spatiotemporal series, and ends up in self-contradiction.

Objection no. 2: an empirical whole in space-time is not intrinsically impossible. Allison provides a subtly different resolution to the antinomies from the TR-P_2

[396] *KdpV* 5:4.

[397] Allison says that Kant's views on space and time are not alternative ontologies, but alternatives to ontology. Henry Allison (2004): 98.

combination of his seminal work, *KTI*, starting in an early paper but also evolving in a major work on Kant's theory of freedom. This resolution consists in the claim that a complete whole, or absolute totality of the world reaching to some unconditioned limit, is not a possible empirical concept. Empirical concepts must be able to refer to an object of possible experience. But, the idea of an absolute whole violates the conditions of possible experience, since we can experience the world only partially and not simultaneously as a whole. Allison applies a similar treatment to the notion of an absolute and unconditioned beginning of a causal sequence. If this is to be an empirical concept, it must refer to an object of possible experience. But, objects of possible experience must be subject to a causal law, and so an unconditioned beginning cannot be an empirical concept.[398]

This treatment of the antinomial conflict is only superficially different from the more recent 2004 treatment involving the notion of transcendental illusion. The idea is that we cannot posit the totality of the world as something really given, because this would transcend the limitations of our anthropocentric perspective, i.e. the limitations of our empirical concepts. We are here again working only within epistemology: the cosmological positions in the antinomies are *epistemologically* flawed, because they posit the idea of the world constitutively, as something really given, in spite of the fact that an idea of the whole is not a proper object of cognition for us. What I am arguing for, though, is that the antinomies are *metaphysically* flawed, insofar as they force reason to posit the unconditioned in a spatiotemporal domain that is pervasively conditioned, and so cannot possibly serve as the arena for reason's unconditioned.

[398] Henry Allison (1976): 243-5.

In *Kant's Theory of Freedom*, Allison makes the same point, that the idea of a spatiotemporal whole, or totality, as either finite or infinite, cannot be an empirical concept, because the complete set of spatiotemporal items is not a possible object of experience.[399] The metaphilosophical therapy of TI, then, is to refrain from the dogmatic and absolutizing tendencies of reason, which seeks cognition of the world as a totality when such a totality is not even a possible empirical concept. The antinomies are dissolved as reason accepts metaphysical totalities as merely regulative standpoints.[400]

The problem with this argument is that it conflates conditions of meaningfulness, and so existence-conditions, with conditions of cognition. The idea of a whole in space-time is meaningful, even though we may not be able to have cognition of it is as whole in a possible experience. A higher intellect, capable of intuition independently of the limitations of our sensibility, could grasp the world as a whole.[401] If the empirical whole is meaningful, then the world can still be, paradoxically, both finite and infinite, and its law of causality would still crowd out reason's demand for explanatory completeness in the form of absolute freedom. Again, the mere epistemic restriction to what is cognizable for us—i.e. what is capturable within our empirical concepts—does nothing to protect the possibility of absolute freedom as a real object.[402] We may not be able to cognize the totality of the world in a single experience, but its existence in space and time is still

[399] Henry Allison (1990): 24.
[400] The infinitistic standpoint of the antithesis of the first antinomy, for instance, is reduced to the regulative demand to always seek more conditions, rather than a dogmatic claim to knowledge of an infinite world. Allison (1990): 24.
[401] An intuitive intellect could produce the world, and so cognize it as a whole. The world would not be given and synthesized piecemeal by an intuitive intellect. *KdrV* B145.
[402] Karl Ameriks notes that the idea that there is some intrinsic impossibility of there being a spatiotemporal whole in experience presupposes an arbitrary restriction on experience. Karl Ameriks (2004): 75.

meaningful, and this would lead to the paradox of an absolutized spatiotemporal framework that is both finite and infinite. Similarly, it is still meaningful for the causal series to exist, even if it cannot be an empirical concept for us, and this would preclude absolute freedom as a real unconditioned object.

What is needed to refute the possibility of an empirical whole is a recognition of a metaphysical idealism about the nature of space and time. If space and time pertain to the subject only, and do not characterize things in themselves, then a whole cannot exist in space-time. The existence of a whole in space-time is not intrinsically incoherent, then, but only on a certain understanding of space-time—namely, one that assigns to space-time a metaphysical idealism. On this understanding, we cannot ascribe the unconditioned properties of a really given totality to space-time, because such an ascription leads to contradictions, and so spatiotemporal appearances can only be conditioned entities that do not partake of reality qua totality. Allison's solution does not truly negate the possibility of an empirical whole; it only negates the possibility of cognition of it. Without negating the possibility of a spatiotemporal whole—which is the negation of the neglected alternative that space and time characterize not only our a priori structures but also the unconditioned totality of reality in itself—reason is not able to secure its aspiration to the unconditioned as a real object, since the spatiotemporal structure of experience threatens to smother the unconditioned.

The mere epistemological modesty of the Transcendental Analytic, which avoids the Leibnizian use of pure concepts to gain rationalistic insight into both appearances and things in themselves (i.e. intellectualize phenomena, so that reflection on pure concepts

gives a complete account of the empirical and intelligible worlds[403]), may resolve the antinomial conflict, but only by sacrificing a metaphysical realism about cosmological objects. The existence of the world as something possibly really given, as well the real givenness of absolute freedom, is jeopardized by paradox even if we recognize that we cannot have metaphysical knowledge of cosmological question because of the limitations of our empirical concepts. What is needed to secure the possible real givenness of absolute freedom, and the non-paradoxical existence of the world, is a rejection of the neglected alternative, which requires metaphysical commitments. We cannot reject the neglected alternative through a mere recognition of the epistemic necessity of space and time. What is a subjective form may also be an objective form.[404]

Objection 3: The resolution of epistemological TI fails to protect reason as a reality-tracking faculty. Since epistemological TI only recommends an epistemological modesty that withholds real givenness to cosmological positions on the world-whole, it does not protect the possible existence of the unconditioned as a real object. The world as an unconditioned totality would have to exist as both finite and infinite, if we have not rejected the objective spatiotemporality in the metaphysics of the neglected alternative. Absolute freedom would have to exist in an objective temporality that subjects each event to a prior cause. Reason's aspiration to the unconditioned is not arbitrary or optional. Kant writes, "Yet the idea of this completeness [i.e. the unconditioned] still lies in reason, irrespective of the possibility or impossibility of connecting empirical concepts to it adequately."[405] But, if reason has to exercise its orientation to the unconditioned only as a

[403] Kant uses this phrase in *KdrV* A271/B327.
[404] This is the phrasing of Trendelenburg, as cited in Allison (1976): 313.
[405] *KdrV* A417/B445.

regulative ideal to avoid contradiction, it takes on a new role as merely establishing formal frameworks without ontological content. I will question in this section whether this non-assertoric perspectival role for reason actually provides reason a genuine satisfaction of its need for the unconditioned, i.e. idea of completeness.

A reason that posits absolute freedom as a mere framework for deliberation, which ends up in contradiction when hypostasized as a real object, is a curious kind of faculty that can be interestingly evaluated in light of certain controversies Kant had with his contemporaries. F.H. Jacobi, for instance, claimed that an excessive trust in Enlightenment rationalism would lead to a Spinozistic denial of human agency, and we need, instead of following reason to its logical conclusions, to embrace a reason-transcending faith.[406] While Allison is committed the rational necessity of the idea of freedom—he notes that we cannot regard ourselves as rational agents without adopting the idea of freedom[407]—the fact that the real givenness of this idea is not secured as a possibility through a metaphysical rejection of the objective structure of space-time makes the rational necessity of this idea suspect. A mere deliberative standpoint involving the idea of practical spontaneity, that one cannot guarantee has a real realization in the world, appears more akin to a faith than something rationally incorporated into a system. The specter of the neglected alternative, which would preclude absolute freedom, means that one's idea of absolute freedom can never be systematically integrated, as a real object, into one's empirical cognitions. Allison's

[406] This is the premise of Kant's brief but important essay, "What Does It Mean to Orient Oneself in Thinking?" in *R*: xxxviii. For a discussion of Jacobi's general suspicion of over-intellectualization, and how the doctrine that nothing comes from nothing inevitably leads to a rejection of a creation theory by a transcendent God, see Dieter Henrich (2003): 98.
[407] Allison refers to the language in *Gr* 4:448.

deliberative standpoint, rather than a rationally integrated and really possible object, appears more akin to a Strawsonian standpoint that constitutes moral responsibility strictly in terms of mere reactive attitudes independent of any theoretical justification.[408]

The orientation of reason means—as Kant explains in a brief essays responding to a controversy between Jacobi and Mendelssohn—that reason, in responding to its own need to seek the unconditioned apart from experiential confirmation, must ensure that its notion of the unconditioned is free of contradiction. Having assessed the unconditioned as free of contradiction, reason can at least coherently think the unconditioned in relation to the cognized objects of the understanding in experience, even though it cannot have cognition of the unconditioned. A reason that fails to establish its aspiration to the unconditioned as free of contradiction is not properly oriented, but indulges in enthusiasm.[409] Enthusiasm is the path of a figure like Jacobi: we can just blindly believe in our absolute freedom, through a "sense of truth," even though rational thought about this leads to contradictions.[410]

Reason's aspiration towards the unconditioned is a need of reason, and it is important to establish the right of reason to aspire towards the unconditioned. The question of whether or not we can actually cognize the object of the unconditioned is separate from the right of reason to think it. Here, we are reminded of the *quid juris* question of the Transcendental Deduction[411]: that is, we are not concerned with the fact that reason thinks the unconditioned (*quid facti*) but with what right (*quid juris*) reason

[408] See the discussion in Gary Watson (1993): 122.
[409] Ibid., 8:137.
[410] 8:134.
[411] For the distinction between *quid juris* and *quid facti*, see *KdrV* A84.

thinks it. The cosmological ideas of the antinomies are not objects of wild speculation. Kant is careful to distinguish the unconditioned from a needless multiplication of causes by which reason only confuses itself. For instance, reason might posit "spiritual beings in nature," but this would only confuse reason because we can identify many causal mechanisms that are grounded in experience.

Kant says that reason "will be satisfied" in its quest for the unconditioned, even if it is prevented from cognition by human epistemic limitations.[412] If reason cannot identify sensible intuitions by which to have objective validation of its ideas, it must satisfy itself by establishing the supersensible unconditioned as free of contradiction, and by gaining some non-contradictory grasp of a possible relationship of the supersensible to the objects of the understanding. Kant rejects the move of Jacobi, in response to reason's annihilation of the possibility of freedom through taking the principle that *nihilo nihil fit* (nothing comes from nothing) to its logical conclusions. Jacobi abandons reason for faith in order to preserve human autonomy.

It is legitimate to question to what extent the Allisonian version of freedom as a regulative standpoint, that cannot be realized as a fact of the matter, differs from the enthusiasm, i.e. non-systematic abandonment of reason, of Jacobi. One might object, as Allison does, that this regulative standpoint is rationally necessitated as a presupposition of one's practical agency, and so it is not a mere article of faith.[413] However, if there is a specter of the neglected alternative, then we can only be rationally necessitated to adopt

[412] *WhdO* 8:137.

[413] Allison is careful to distinguish the phenomenon of acting under the idea of freedom from the optional heuristic of the intentional stance of Dennett, in Henry Allison (2012): 41. The idea of freedom is a rational necessity for the agent, and not a mere stance one adopts according to pragmatic considerations to better predict behavior.

the idea of freedom as a non-assertoric standpoint. It is doubtful that such a standpoint really allows reason to satisfy its need, since one might argue that such satisfaction requires the real possibility of objective realization of absolute freedom. Can the satisfaction of reason lie in an idea of completeness that is fictional, a merely formal framework not grounded in reality? It is doubtful that something can be both rationally necessary and also not track reality. Dana Nelkin, for example, notes that, if an event is determined, it is determined regardless of how we may describe it. If the neglected alternative is true, and temporality, with its causal law connecting us to determining grounds, characterizes us objectively, then we are determined even if we abstract from epistemic conditions to think of ourselves as free. I am not rationally justified in thinking myself as free through a special mode of description, if it has been metaphysically established that I am determined.[414] Kant, similarly, claims that, though thinking regulatively is mere "relative supposition" that does not involve absolute knowledge of the thing in itself, the relative supposition of a regulative framework must at least correspond to an indeterminate object, a "Something in general."[415] What is rationally necessary cannot be mere mode of description that can be satisfactorily excluded from one's view of the world, if one can account completely for the person in terms of a non-agential mechanistic perspective. Explaining action in terms of a deliberative perspective is not a genuine explanation if deliberation is not guaranteed a metaphysical realization and one's conditioned status in space-time controls one's behavior.

[414] Dana Nelkin (2000): 565.
[415] See the distinction between *suppositio relativa* and *suppositio absoluta* in *KdrV* A678/B706.

Since the empirical series might block absolute freedom in epistemological TI, any belief in absolute freedom would be pathological in the Kantian sense. Kant uses the term "pathological" to refer to love that arises from feeling, and not from the will.[416] In the context of belief in absolute freedom, it is pathological when it based on a mere feeling, rather than rational necessity. Epistemological TI leads to a pathological belief in absolute freedom, based on a mere private opinion that the empirical series is not absolute. We can believe in freedom as a real object—in response to the question that continually haunts us, according to Allison, are we really free?[417]—only because we may want to, not because there is any rational guarantee that an objective temporality does not block the possibility of absolute freedom. Since both the mechanistic and agential standpoints both cover all the facts of the matter, which one is adopted depends on the preference of a given moment. Again, referring to Nelkin, it is rational to maintain belief in universal determinism, while having the desire that one is free. It is by downgrading one's stance towards freedom to a desire that one is able to avoid the contradiction between the belief that one is free and the belief that one is determined.[418] But, downgrading the practical cognition of oneself as free to a mere desire (which Allison does not do—I only use this as an analogy) fails to take into account the important contributions of practical reason, in which Kant retains his credentials as a rationalist metaphysician in the Leibnizian mold. Practical reason is, indeed, supposed to have

[416] *GMS* 4:399.
[417] Allison refuses to give an answer to this question, claiming that it is a symptom of the transcendental illusion, in Allison (2004): 49.
[418] Dana Nelkin (2000): 568.

primacy, such that theoretical reason must accept, as a legitimate increment added to itself, the findings of practical cognition.[419]

For epistemological TI the idea of freedom is a regulative stance that is also rationally necessary, not a mere desire. But, there is a kind of doublespeak here: belief in freedom in epistemological TI has to be rationally necessary as a standpoint, but cannot be rationally necessary as something really given. Freedom as a merely non-propositional, fictional standpoint cannot provide the kind of systematized completion of metaphysics for which Kant was looking.[420] It may be the case that deliberation without the idea of freedom is a contradiction. However, failing to guarantee the metaphysical basis of the idea of freedom also presents a contradiction with deliberation. Dana Nelkin's query, of whether belief from a standpoint is even an intelligible idea, is a legitimate one. If we only believe in freedom as a non-assertoric stance in certain contexts, we do not really believe it, since belief makes a commitment to the nature of reality *simplicter*.[421]

Typically, when one discusses beliefs from different standpoints, one is expressing an indecision, such that one recognizes considerations favoring different beliefs that prevent one from committing to belief in either position.[422] But, this sort of indecisive agnosticism is not what Allison has in mind. In Allison, the practical standpoint under the idea of freedom is rationally necessary, even though we cannot

---

[419] *KdpV* 5:121. See Ostaric (2017): 110.
[420] I am in agreement here with Anja Jauernig and her judgment that Kant did not entirely abandon Leibnizian rationalist metaphysics, although he took a roundabout route through practical reason. Anja Jauernig (2008): 52. For another confirming view, see Ameriks (2003): 134.
[421] Dana Nelkin: (2000): 569.
[422] See Nelkin's example of someone who, from the standpoint of being an American, believes *p*, but from the standpoint of being a Christian, believes not-*p*. Ibid., 569.

guarantee the possibility that the idea of freedom has an objective existence.[423] Taking a practical standpoint, when the causal series is not appropriately qualified to protect the idea of freedom, is what Kant would call an enthusiasm—a cognitively empty belief not properly vetted for contradiction—and would not figure in an integrated rational system.[424] If, as Allison claims, we are free only from a standpoint, belief in freedom cannot be rationally necessary, since rational necessity transcends standpoints. What is rationally necessary cannot be based on a contingent psychological need to take oneself as free, having to do with a non-factual standpoint we abandon when in the third-person perspective. The price Allison pays for agency is the steep one of an antinomial grasp of the world, in which the complete picture of both our agency and our existence in space-time is disunified.[425]

Objection no. 4: epistemological TI and resolutions of sets of antinomies. The mathematical antinomies both deal with the series of conditions within appearances, and are properly resolved by claiming that both the thesis and the antithesis arguments are false.[426] The dynamical antinomies conjoin an intelligible condition to the empirical series. The resolution of the dynamical antinomies holds that both the thesis and the antithesis may be true.

This distinction is of interest because it does not seem as though the diagnosis of epistemological TI, with respect to the errors of the Antinomies—i.e. that they have to do

[423] Allison therefore resembles a Humean compatibilist that insists on the necessity of approbation and blame even though we might believe in divine determinism. Discussed in Pereboom (2001): 91.
[424] Guyer makes a similar point about the principle of systematicity: the mere postulation of systematicity, apart from evidence for it independent of one's wishes, does not seem rational. Guyer (1979): 50.
[425] Galen Strawson discusses a lack of unity in theoretical truth in Strawson (1986): 277.
[426] *KdrV* A530/B558.

with an undue epistemic ambition to have cognition of the *an sich* reality of things—can ground different resolutions of the two sets of antinomies. Both the thesis and the antithesis of all antinomies hypostasize the unconditioned in taking a stand on some cosmological question. Once we undergo the therapy of TI, in Allison's version of it, we acquire a new mind of epistemological modesty, in which we refrain from taking a God's eye view of things. It seems TI as epistemological modesty—i.e. an anthropological rather than a theocentric perspective—would yield a blanket resolution of all four antinomies: contradictions would be dissolved in the name of a modesty that repudiates all thesis and antithesis arguments as examples of epistemic overreach into *an sich* reality. A blanket agnosticism, however, does not appear to provide a ground for different truth-values for mathematical and dynamical antinomies.

Again, it is the specter of the neglected alternative that complicates the different resolution strategies in the dynamical and mathematical antinomies. When we deny the neglected alternative, we know that the world of spatiotemporal appearances cannot be a totality, because space and time are ontologically dependent on the subject and so not capable of absolutization as either a finite or infinite whole. If we do *not* deny the neglected alternative, it would still be possible that the world in space-time is a whole, since an objectively existing spatiotemporal structure of reality would be subject to absolutization as an unconditioned totality. We would have to accept the possibility that both the thesis and the antithesis may be true, and that what appears paradoxical to our reason—i.e. a finite and infinite world—might actually exist. To directly address Allison, without a rejection of the neglected alternative, we can still infer, from the fact that world

in space and time is not finite[427], it may be infinite, and vice versa, since, given the neglected alternative, the unconditioned totality may take place in space-time. If we reject the neglected alternative, we also know that the causal series in time is not an objective structure of reality in itself and so is not amenable to absolutization. It therefore may be the case that a faculty of absolute freedom may be true. But, if we do *not* reject the neglected alternative, we cannot guarantee the possibility that causality in time is not an absolute structure that crowds out absolute freedom.[428] We cannot guarantee, therefore, that it *may be true* that we have absolute freedom. The metaphysical rejection of the neglected alternative allows for the modal distinctions that Kant provides with respect to the resolution of the different sets of antinomies.

Kant's Commitment to the Metaphysical Contingency of Space-Time

In this section, I will show how the metaphysical idealism about space-time, by which it does not track the absolute structure of reality *an sich*, is not an aberrant view, but actually one that has roots in Kant's early scientific and philosophical work. The idealistic trajectory of Kant's pre-Critical work will lend credence to the idea Kant is invested in a certain ontology of space and time, and is not committed to them solely as

[427] Allison claims that, in the first antinomy, reason mistakenly assumes, because of transcendental realism, that what is not finite must be infinite. Allison (2004): 387.
[428] Kant is clear that the causal law of temporal succession is a condition of apprehension, not of being. See *R* 5263, cited in Guyer (1987): 65.

epistemic conditions, as Allison claims. I have shown how this ontology of space-time is essential to securing the possibility of the unconditioned as a real object.

Non-fundamentality of space-time in True Estimation of Living Forces. In a very early pre-Critical scientific work, *True Estimation of Living Forces*, Kant promotes the idea that it is the force of interaction between substances that precedes and grounds spatial location. Force provides connection between things, connection provides order, and order is finally the basis for space.[429] Forces are more fundamental than space, and so spatiality is not an absolute and irreducible substance in Kant's earliest works. Kant, in this early work, criticizes those who reduce forces to their empirically observable effects in space-time.[430] Kant was critical of the formidable consensus of Newton, Locke, and Descartes that physical extension is an irreducible ontological primitive of bodies. Kant accepts the metaphysical-dynamical notion of matter, which establishes that force is necessary for matter to fill a space, and so have extension in the first place.[431] Kant explains how a metaphysics of forces, rather than one of mere extended bodies in motion, can explain how the soul, a non-extended substance, is able to act on the body.[432] Kant of the *True Estimation* acknowledges the possibility of the soul's action on the body—i.e. an intelligible force acting on a spatiotemporally extended substance—by denying that the only power in the world consists in extended bodies in motion.[433] With the idea of a force that is not identifiable exclusively with extended matter in motion, Kant is able to

[429] *IK* 1:23.
[430] See Eric Watkins (2005): 271.
[431] See Eric Watkins (2001): 114.
[432] *IK* 1:21.
[433] See the discussion of *IK* in Eric Watkins (2005): 102-104.

bridge the gap between soul and body that rejects the causal isolation we find in Leibnizian pre-established harmony.

Kant is clearly more Leibnizian than Cartesian with respect to the famous *vis viva* debate of the 17th century, in which two paradigms of natural philosophy were adjudicated: one in which spatial extension is fundamental to matter, and one in which a non-spatial forces are fundamental to matter.[434] The Leibnizian conception of matter incorporated the Aristotelian commitment to substantial forms, in virtue of which a force acts as a sort of underlying blueprint for the development of matter.[435] This view clearly presages Kant's view of the intelligible character in the first *Critique*, in which something more fundamental than empirical events in space-time provides a structure to them. Kant throughout his career endorses a dynamical theory of matter: matter cannot have a stable shape but for attractive and repulsive forces. Attractive forces keep matter from dispersing, while repulsive forces keep it from collapsing.[436]

*Non-fundamentality of space-time in New Elucidation***.** In *New Elucidation*, Kant claims that, without a connection between substances, succession and time itself would disappear.[437] This is the Principle of Succession, which indicates that no change can happen to a substance unless they are already in connection with other substances. Substances in isolation are immutable in themselves and cannot be the basis of succession, and so an activity connecting substances, and not their mere existence, is the

[434] See the introduction by Eric Watkins in *Natural Sciences* (2012): 3.
[435] The idea of an entelechy is a goal within an object, akin to computer software of DNA, that controls the development of matter. Ibid., 688, note 24. Kant was aware of the objection of the empiricists that the notion of an entelechy, and metaphysical forces underlying empirical extension, is obscure. Eric Watkins (2001): 111.
[436] See Rae Langton (1998): 125.
[437] *ND* 1:410.

necessary basis of change, even time. There are clear similarities between the Principle of Succession and the Third Analogy of the first *Critique*, in which Kant claims that objective simultaneity of events requires reciprocal dependency of substances.[438]

*Non-fundamentality of space-time in the Inaugural Dissertation***.** Kant makes clear commitments to the subjectivity of space and time in his *Inaugural Dissertation*. Here, his view shifts from one in which space and time depend on metaphysical forces to one in which space and time are irreducible features of subjectivity, but the contingency and non-fundamentality of space-time remains. The view of the empiricists, that spatial extension is an ontological primitive, remains something that Kant rejects, as well as the Newtonian view of time, in which we can distinguish between relative time, which characterizes sensible objects, and absolute time, which is more fundamental than a relationship to sensible objects.[439] Time is not some absolute container more fundamental than things for Kant, but a structure of sensibility in which we coordinate all appearances. The idea, moreover, that space is an absolute receptacle of things—something metaphysically fundamental—is, according to Kant, is an empty fabrication of reason that posits infinite relations without any objects to enter into relationships.[440]

The subreptic axioms teach us that we must not take the conditions for sensible representation as conditions for the thought, through the understanding, of things in general. The thought of a thing as such does not have to include the conditions necessary

[438] One can find the Third Analogy at *KdrV* A211/B257.
[439] See the discussion in the introduction to the *Leibniz-Clarke Correspondence,* by H.G. Alexander (1956): xxxiv.
[440] Kant's views on space and time come in sections 14 and 15 of the *Diss*. See Guyer's discission at Guyer (1987): 350.

for its sensible representation.[441] This is exactly the kind of separation between reason's thought of the unconditioned, and the spatiotemporal series of empirical cognition, that is necessary to resolve the antinomial conflicts. If space and time are ontologically fundamental, there would be no problem in conflating the conditions of representation of a thing, and the conditions for its existence.[442]

Some of the problems of taking space and time to be ontological conditions of the existence of things have to do with trying to find a location for an immaterial substance like a soul. If space and time are ontological primitives, then everything, even a soul, must exist at some place and time. Trying to locate God, also, at a particular time, leads to the fear that God's decision to create the world at a certain time, and not another, was arbitrary and so imperfect.[443] Kant endorses Leonhard Euler, in this discussion of the virtual, and not the local presence of the soul, as extremely perspicacious.[444] A virtual presence is an action that acts at a certain location, but does not itself have a location. Euler defines spirits as entities that lack extension, and debunks the Wolffian theory that spirits are therefore comparable to mathematical points. The Wolffian notion of a spirit has a location in space-time, insofar as it the smallest possible particle of matter that occupies an indivisible geometrical point. However, for Euler and Kant, questions of extension are inapplicable to spirits, just as they are inapplicable to stretches of time (there cannot be, for instance, a meter of time). Instead of occupying a particular place,

[441] Discussion of the subreptic axioms comes in section 26 of the *Diss*.
[442] Space and time for Kant are not ontologically fundamental, because they constitute the form of only the sensible world. In the argument for the First Analogy, Kant states that time is the permanent form of inner *intuition* (*KdrV* A182/B225). In the *Prolegomena*, Kant states that space is the fundamental framework, or substrate, of *intuitions* of particular objects (*P* 4:322). The theme here is that space and time are fundamental to intuitions, not necessarily to things in themselves.
[443] *Diss*, 2:414.
[444] Ibid.,

spirit, just like metaphysical forces, act at certain locations while themselves lacking a location.[445]

The Non-Fundamentality of Space-Time and the Metaphysics of Agency

Epistemological TI never establishes the non-fundamentality of the empirical series. Spatial extension and temporal duration may be ontologically fundamental while at the same time representing the necessary and universal structure of experience. I have just identified a whole history in Kant of dealing with issues in physics and metaphysics having to do with the metaphysical non-fundamentality of space and time.

Failure to establish space-time as metaphysically non-fundamental means that we can only establish transcendental freedom as a perspective, not as an ontological fact of spontaneous causality. In this section, I will apply the non-fundamentality of space-time to a metaphysics of agency. Having established that a metaphysics of space and time is central to resolving the anitnomial conflicts, the question remains of how exactly agency would work given this metaphysical framework. First, I will deal with the objection that positing an underlying non-spatiotemporal domain of unconditioned freedom makes the empirical series illusory. Second, I will raise the objection that freedom from a standpoint is causally inert. Third, I will explain how the relationship between categorical and

[445] See Euler's Letter XCII in "Letters to a German Princess," in *Kant's Critique of Pure Reason: Background Source Materials*, ed. and trans. by Eric Watkins, (Cambridge: Cambridge University Press, 2009). Crusius is another thinker who thinks that space and time are ontological principles of existence. See Watkins (2005): 82.

dispositional properties can show the causal relevance absolute freedom in a metaphysical TI in a way that avoids problems with event-causal accounts of freedom.

*Non-Spatiotemporal unconditioned properties compromise empirical realism***:** Allison claims that there is no context-independent fact about our freedom, since any kind of ontology underlying empirical appearances would relativize them and compromise empirical realism.[446] That is, if there were some capacity for absolute freedom underlying the appearance of ourselves as determined, the appearance of determinism would be illusory.

But, Karl Ameriks debunks this so-called "demoting thesis."[447] Ameriks most compelling analogy has to do with the appearance of jaundice. The appearance of jaundice indicates something about the nature of the person in herself, i.e. some underlying pathology. Let us say that a new biological phenomenon occurs in which jaundice appears in healthy people. Jaundice would now no longer indicate anything about people in themselves. Just because, however, jaundice does not correspond to the *an sich* reality of the person, does not make it illusory. It is still a real appearance. In the same way, something really appears spatiotemporal, even though it may not correspond to the reality of things in themselves.[448]

The causal inertness of freedom as a perspective. Freedom from a practical perspective, that does not correspond to a real ontological power, would be causally inert. It is widely accepted that genuine causation is supposed to instantiate a causal law—if A

[446] Henry Allison (2004): 48.
[447] See Karl Ameriks (2012): 93.
[448] This argument is found in Karl Ameriks (2004): 36.

causes B, it has to be able to do this in a lawful way that is repeatable, otherwise the relationship is merely accidental.[449] But, if freedom is a mere perspective, and the only genuinely causal and lawful connections pertain at the empirical level between conditioned events, freedom is a mere meaningless dangler. Without a genuine non-empirical form of causality, freedom qua perspective is completely limited by the causal relationships at the empirical level. The kind of decision the agent takes under the idea of freedom in the agential perspective actually has no relevance to what the agent actually does. Even if we take a singularist view of causation, that views it as realizing only an instantaneous law for a given moment[450], freedom qua perspective is completely constrained by causal relationships at the empirical level. If we adopt the position that the causal series in space-time may be absolute, then the only way the practical perspective can have any relevance to this world is through entering into the already established causal relationships in the empirical series.[451] Given that natural laws at the empirical level sets the parameters of a perspectival freedom that cannot act as a genuine contra-causal power, the only way that perspectival freedom can be the genuine source of action is in virtue of certain unconvincing compatibilist versions of authorship. For instance, freedom qua perspective could only be the "author" of its actions insofar as it imparts

[449] Jaegwon Kim notes that the nomological view of causation is the received view in Jaegwon Kim (2002): 172.

[450] Henry Allison mentions the idea of instantaneous laws in rebutting the idea of Friedman that a causal law linking events guarantees continuity beyond a single instantiation of a causal relationship. Allison (1993): 298.

[451] This criticism echoes Kim's criticism of anomalous monism: the mental perspective, since it is completely constrained by the laws at the physical level, is merely epiphenomenal. Kim (2002): 172.

interpretative significance, while being completely dominated by lawfulness at the empirical level.[452]

Allison's reductive understanding of the order of ends framed by practical reason means its idea of freedom is a merely regulative projection.[453] The idea of freedom qua regulative projection becomes a dispensable postulate that lacks any genuine explanatory role with respect to the agent's behavior. If freedom is a merely regulative perspective, it is hard to see how it is not merely driven by external impulsion, rather than being the author of its own behavior.[454] If the Incorporation Thesis is a merely conceptual claim, and not a metaphysical one about what really takes place[455], it is hard to assign to it any integrated role in agency at all. The Incorporation Thesis becomes a fiction that one thinks of oneself in a way that one is not, as if something earthbound were to think of itself as flying. If the idea of freedom is not allowed a possible metaphysical realization, it is not clear how Allisonian freedom is anything more than Lockean "thinking extended matter," i.e. an entity entirely controlled by material causality that nevertheless has certain capacities for deliberative self-reflection.[456]

Categorical and dispositional properties and the causal relevance of freedom in metaphysical TI. If we limit the empirical causal series, in virtue of a metaphysical idealism, so as to make possible the idea of an unconditioned causality at a metaphysically basic level, we do not need to condemn practical agency to causal

[452] Michael McKenna discusses this theory involving interpretative significance of action in McKenna (2008): 201.
[453] Allison uses this language in Allison (1990): 38-9.
[454] Kant rejects the idea that reason is driven by mere impulsion at *GMS* 4:448.
[455] Allison makes this admission in Allison (2008): 453.
[456] Locke considers the idea that senseless matter could take on capacity for thought in Locke (1996): IV.iii.6.

irrelevance.[457] Contemporary metaphysics can articulate the difference between the intelligible and empirical characters in terms of categorical and dispositional properties. The empirical character represents, one might argue, a dispositional property, insofar as it arises in relation to interaction with the spatiotemporal form of human sensibility. A dispositional property has to do with a potentiality of an object activated in certain circumstances. Fragility is a dispositional property insofar as it is manifests itself when an object is placed under pressure.[458] Dispositions are not identifiable with specific occurrences, but stable tendencies: an object has the dispositional property of fragility even when not subjected to pressure. The stability of dispositions, then, involves not only events, but grounds.[459] The incorporation of grounds into a metaphysics of agency rescues it from problems associated event-causal metaphysics. If an agent's exertion is a mere event, and not based in a ground that abides across time, it becomes difficult to differentiate the event from a chance occurrence that lacks roots in the agent's character.

We can establish categorical properties as those properties something has unconditionally, independently of any relationship. The intelligible character, which the person has independently of human sensibility, can be understood as a categorical property. The fact that the way something is categorically grounds the way it is dispositionally leads us out of the causal isolation of free will as a mere practical standpoint. The intelligible character represents the unconditioned nature of a person—its

[457] Allison states that Kant's goal in judging the relationship between nature and freedom is to articulate a relationship of compatibility without reducibility. If we do not establish that the empirical domain is not metaphysically absolute, the only way we can preserve compatibility with freedom is through reducing it to a merely causally irrelevant perspective. Allison (2001): 202.

[458] Stephen Mumford (1998): 20.

[459] A model of causality that appeals to grounds is therefore very different from a Humean event-event model of causality, in which individual events do not endure over time but simply pop into and out of existence. See the discussion in Eric Watkins (2005): 245.

categorical properties—that ground—and so is not causally isolated from—the dispositional properties of the empirical character. For example, the fact that sugar is soluble in water—a dispositional property having to do with a certain conditional relationship—is based on unconditional properties having to with the chemical nature of sugar in itself.

We can further appeal to the notion of maxims to understand how the categorical structure of the agent shapes the dispositional properties it has in relation to sensibility. A maxim is a subjective principle that mediates between the moral law, which is the objective law for all rational beings, and an individual's private psychological drives.[460] Onora O'Neil has explained how maxims function as umbrella principles that orchestrate various specific intentions. For instance, the umbrella principle of hospitality towards guests might prompt the specific intentions of making tea, offering a chair, smiling, etc.[461] Maxims, as umbrella principles, represent stable grounds of action underlying specific events—i.e. they act as dispositional properties.[462]

Maxims provide the point of entry into mapping the model of causality involving temporally indeterminate grounds underlying events, articulated in terms of categorical and dispositional properties, on to a metaphysics of agency. In the time-series of the empirical level, agency manifests itself in the form of temporally specific events. Underlying these individual actions are certain umbrella maxims that are temporally indeterminate. The maxims correspond to the categorical structure of the agent that is

[460] See the definition of maxim in *GMS* 4:421. Also, 4:402.
[461] Kant makes it clear that maxims are broad principles that permit a certain latitude (i.e. playroom) in what way one follows them, in *MdS* 6:390.
[462] Onora O'Neill (1983): 394-7.

unaffiliated with a spatiotemporal sensibility, since the maxims do not manifest themselves as specific events. Just as categorical structures ground dispositional structures, the structure of maxims at the intelligible level shape the way events unfold at the empirical level. A person is *disposed* to act in certain ways in virtue of underlying maxims. The maxims structure the person in a middle domain between potentiality and actuality. An individual with a maxim has a certain dispositional characteristic even when not in a situation that would activate the disposition.

This metaphysical structure of categorical and dispositional properties abides across time in the form of implicit potentialities underlying actions that appear at a specific time. This temporally indeterminate metaphysical structure provides a way to respond to objections to certain theories of free will that claim that actions in the theory are driven by luck, and so are not free. For instance, even agent-causal views, that ascribe to agents the capacity to act in different ways given the same past and laws of nature, have been attacked as basing action on mere luck, since the identity of the past and laws of nature appears to mean that the agent's decision in different scenarios is arbitrary, i.e. has no explanatory grounds.[463] The assumption here is that the only way an action can be non-arbitrary is through some difference in the temporal series of events leading up to the action. However, the incorporation of categorical and dispositional properties, in the form of a non-spatiotemporal metaphysics that does not manifest itself in the form of events, allows us to articulate a way in which an agent might act in different ways, even if all the events leading up to her action are identical. An agent acts differently, in spite of an

[463] See Meghan Griffith's discussion in Meghan Griffith (2010): 44. Alfred Mele makes this point also in Mele (2006): 8.

identical past, because of different categorical and dispositional properties that do not manifest themselves in the form of events. The agent's act does not *happen to* her—it is something she *does*[464]—because, even though the series of past events and the laws of nature are identical, there is a certain underlying metaphysical structure having to do with properties that do not express themselves in the form of events, that shape the way an agent acts. The agent causes her action *as a substance*, and not through events. In general, an underlying metaphysical structure of temporally indeterminate grounds, that complements events, gives the agent a way to gain ownership over her own psychological events that would otherwise occur through chance, or through a deterministic series external to the agent.[465] The causal relevance of the metaphysics of the agent's intelligible character not only meaningfully integrates it with the empirical character, but also avoids certain issues having to do with a lack of agential control.

Conclusion

In this chapter, I explained two competing attempts to resolve the antinomies in Kant's first *Critique*. On the one hand, there is the resolution of epistemological TI, which focuses on walking back a transcendental realism which absolutizes the cosmological ideas. Metaphysical TI embraces a metaphysical idealism about the empirical series. I showed how epistemological TI does not really resolve the antinomies

[464] See Griffith's distinction between doing and happening in Griffith (2010): 46.
[465] See the luck objections to agent-causation in Derk Pereboom and Michael McKenna (2016): 244.

in a way that allows reason to secure the possible existence of its idea of the unconditioned. I then developed the historical theme in Kant's work of a commitment to the metaphysical non-fundamentality of space-time, to show that there is an oversight in the reduction of TI to an epistemological theory. Finally, I explained the deficiencies in a metaphysics of agency that reduces agency to a standpoint, which include the causal irrelevance of one's agency and an inability to rescue the agent from acting on mere luck.

It needs to be noted that many have tried to show how belief in both freedom and determinism from different standpoints, i.e. the practical and theoretical, is rational, even without an ontological distinction allowing us to limit the theoretical perspective as ideal.[466] For instance, the theoretical and practical standpoints, one might say, involve different interests, and these interests present different paradigms that allow us to rationally maintain belief in both our freedom and determinism. We might argue that we have an interest in scientific explanation from the theoretical perspective, while we have an interest in moral evaluation from the practical perspective. One might also argue that the justificatory basis of belief in freedom and determinism is different. So, we are not to take belief in freedom as a theoretical proposition about the nature of the world, but rather a presupposition of our practical agency.[467]

[466] Allison's view of freedom in Kant, for instance, is a merely conceptual one, i.e., it involves a practical point of view involving ourselves as rational agents and a theoretical point of view involving ourselves as part of nature. This merely conceptual distinction has nothing to do with a metaphysical distinction involving a causal power outside of nature. See Karl Ameriks (1999): 828.
[467] Dana Nelkin identifies these possible recourses for standpoint theory in Dana Nelkin (2000): 573.

However, these sorts of arguments, which compartmentalize reason into different and isolated domains, fall away in light of Kant's doctrine of the unity of reason.[468] Practical reason for Kant is not some degraded capacity that only responds to our feelings—i.e. it does not merely pathologically respond to our inclinations towards happiness, and posit, for instance, out of sheer desire, "Mohammed's paradise or the fusion with the Deity of the theosophists."[469] Practical reason is a genuinely a priori faculty—i.e. it is not driven by mere inclination—that is unified with theoretical reason. The unity of reason means that theoretical reason must accept the a priori judgments of practical reason. The unity of reason does not compartmentalize different domains of reason according to justification conditions or interests, and it does not relativize certain claims about the world to a merely perspectival status. Instead, Kant integrates reason in its different roles[470], and this requires a certain metaphysics that ensures that theoretical reason's assessment of the self as determined cannot pretend to be an absolute statement on our natures.

[468] This topic is too complex to give full treatment in the conclusion of this paper, so I raise it only in a cautionary way that shifts the burden of proof on those who would defend the idea of practical reason as a non-assertoric standpoint.
[469] See *KdpV* 5:121.
[470] As Lara Ostaric points out, the need of reason for the unconditioned is not merely the need of practical reason, but of reason as such. Lara Ostaric (2009): 157.

CHAPTER 4

THREE CONTEMPORARY LIBERTARIAN THEORIES OF FREE WILL IN LIGHT OF KANTIAN METAPHYSICS OF AGENCY

Contemporary libertarian theories of freedom adopt a fundamentally different kind of metaphysics from the Kantian metaphysics I have been developed in chapters one and two, involving an indeterminate level of spatiotemporal empirical cognition, along with an intelligible level thought through the pure categories that involves a structure of maxims that provide grounds for the of flow of events in the empirical series. In this chapter, I will show how contemporary libertarian theories are fundamentally flawed, making agency indistinguishable from events that occur through luck and providing no basis for sufficient control. My diagnosis of the root problem of contemporary libertarianism is that it adopts an overly rigorous empiricism which seeks to articulate agency in terms of empirical events, and fails to adopt the Kantian program of situating empirical impressions in a rationalist framework of a priori structures of the understanding as well as an immaterialist metaphysics.

In section one, I will give a brief account of the kinds of contemporary libertarianism I will address. In section two, I will consider contemporary libertarianism in terms of to what extent it can articulate a notion of character—what Kant calls *Gesinnung*—which is an underlying nature of goals, character traits, and dispositions that can provide an account of individual choices that roots them in the agent, and goes

beyond the mere happenstance of sequences of events.[471] In section three, I will explain the virtues of a metaphysical account of transcendental idealism for articulating a notion of character, that can account for both the stability of character, which allows us to understand individual choices in terms of fundamental maxims of an individual, as well as an ability to freely invent one's character. In section four, I will consider the issue of world-constitution for the metaphysics of causally discrete events, which informs contemporary libertarian theories, and the Kantian metaphysics in which a non-spatiotemporal form of potential mutual interaction is more fundamental than spatiotemporal events. In section five, I will show the merits a regressive epistemology[472] in terms of its resources for delivering us from the neo-Humean empiricism of contemporary libertarianism and also the merely epistemic spontaneity of the Allisonian agent.

Three Kinds of Contemporary Libertarianism

Event-causal libertarianism. Event-causal libertarianism (which I will hereafter abbreviate as ECL) holds that free actions are caused by events. We can understand events as fleeting and discrete occurrences, which are detached from an interconnecting metaphysical structure (i.e. power), and which accordingly do not exercise any causal

[471] Henry Allison claims that, without a *Gesinnung*, individual choices are reduced to a liberty of indifference, in which a detached will selects between different alternatives solely in virtue of the external pressures that happen to apply to a given moment. Henry Allison (1990): 137.
[472] I derive this terminology from Karl Ameriks's essay on the Transcendental Deduction, in Karl Ameriks (2003): 51.

constraint on events in proximity to them.[473] The thesis of event-causation is an implicit critique of accounts of free action that incorporate substance-causation. Substance-causation connects a substance directly with an event, so that it is not mediated by some event, occurring at a specific time, in virtue of which the substance causes its action. ECL also holds that some indeterminacy prior to the action must characterize the causal sequence of events leading up to action. What occurs in ECL is that there are competing sets of motivations leading up to an action, consisting in different strands of beliefs and desires, and an indeterministic process, internal to the agent[474], in which the agent settles on a particular course of action. Action in ECL differs from non-action in virtue of being mentally caused, by psychological events of beliefs and desires[475], and also in virtue of an indeterminism[476] leading up to action that prohibits a sufficient cause for the action. Key representatives of this view are Mark Balaguer[477] and Robert Kane[478].

*Non-causal libertarianism***.** The problem with ECL is that an ECL agent passively suffers its action to occur to it, since the action is a fallout of different sets of reasons in an indeterministic resolution of a conflicted decision. If there is no sufficient reason for the decision, given the indeterministic process, then there is no agential control, and the

[473] There are other conceptions of events in the philosophical literature. I am referring specifically to Humean events, and since ECL appeals to indeterminism, in which past events do not exercise causal control over the events that follow them, I think Humean events provide a fitting comparison. Eric Watkins, (2005): 17, 230, 232-5.
[474] Internal luck has to do with indeterminism in an agent's psychology, whereas agent-external luck has to do with uncertainty in the agent's outward situation. Mele (2006): 109.
[475] ECL is therefore a causal theory of action, that defines action in terms of causal history. Mele (2006): 113.
[476] Indeterminism involves probabilistic causation, whereas determinism means one unique future given a past and laws of nature. Timpe (2013): 22.
[477] Mark Balaguer (2010): 71.
[478] Robert Kane (1996): 136.

agent operates according to luck.[479] Non-causal libertarianism, since it recognizes the problems associated with tying action to a causal desire-belief sequence, goes to the opposite extreme of severing action from any causal process altogether, instead of incorporating a multi-dimensional metaphysics involving character. Stewart Goetz has developed a version of NCL in which choices are intrinsic features of the mental properties he calls powers. These choices are intrinsic insofar as they arise completely from the exercise of the mental property of a power, and cannot be reduced to some kind of causal sequence attached to the mental property. It is also important to emphasize that these mental powers are properties, and so NCL is clearly at least as hostile to non-empirical metaphysics of substantial agents as ECL, if not more so. Choices in NCL are primitive events, insofar as they are choices entirely in virtue of their mere occurrence, and not in virtue of some complex underlying causal structure.[480] Another major representative of NCL is Carl Ginet.[481]

*O'Connor's naturalistic agent-causation***.** It is not clear how NCL resolves the luck objection that devastates ECL. If ECL is problematic because the competing motivations simply fall out in an indeterministic manner, absent agential control, removing the power of choice from any causal structure altogether just exacerbates the problem. A possible solution to these problems could be an agent-causal account: here, it is not indeterministic or intrinsic mental events that generate/constitute action, but direct causation by the agent qua substance. The agent-cause differs from event-causation

[479] Pereboom makes the objection that ECL provides leeway through indeterminism, but at the expense of control. Pereboom (2001): 55.
[480] Stewart Goetz (2008): 8-9.

[481] Carl Ginet (1990): 13.

insofar as events cause one another as a matter of course—an antecedent event triggers a subsequent event that follows from it according to physical laws. A billiard ball reacting to the collision of another billiard ball does not direct itself, and can only continue the flow of events according to physical laws. An agent-cause responds to external stimuli as an activated power of self-direction, that freely determines itself in a direction of its choice.[482] The direct connection between the agent qua substance, in a way that is irreducible to events[483], prevents us from understanding the agent as a spectator on her action. Key representatives of this view are Roderick Chisholm[484] and Timothy O'Connor.[485]

O'Connor's version of agent-causation takes a curious twist that suggests he is caught between incompatible interests. Events have clusters of properties which consistently regulate the way they respond to different circumstances. The agent-cause has special properties by which it instead is activated so that it is able to self-determine to select certain options at will, in a way not subject to physical laws. O'Connor's embrace of both event-causal powers as well as agent-causal powers forces him to accept both singularist and non-singularist forms of causation. That is, in event-causation, since the properties of a thing dispose it to react in regular ways to some cause acting on it, causation is non-singular—i.e. it has regularity, and so does not just pertain to a single causal relationship. For example, the cluster of properties associated with hydrochloric acid cause it to burn the skin. This is a lawful relationship involving a mechanical transfer

[482] Timothy O'Connor (2000): xiv. This distinction actually corresponds to Crusius's distinction between existential grounds and an active power. See Watkins (2005): 84.
[483] There is some worry that agent-causation is not substantive, i.e. it is possible to reformulate it in terms of events. Dana Nelkin, (2011): 84.
[484] Roderick Chisholm (1976).
[485] Timothy O'Connor (2000).

from circumstances to effects. In the case of agent-causation, on the other hand, causation is singular, insofar as the special properties of the agent allow it to deliberately select among different alternatives in a self-determining way.[486]

There is a metaphysical quandary in the attempt to integrate these two forms of causation. Causal laws and antecedent grounds in time can account for the non-singularist event-causation, in which events simply fall out in a lawful process from the causes that precede them. But, it seems an intelligible level, outside of the empirical series in time, is needed to account for the teleological and singular agent-causation, since it is able to determine itself independently of the empirical series. There has to be an account of how the agent gets the special capacity to determine itself in a way that does not have to track the empirical laws that characterize its context.

The singular causation characteristic of agent-causation is, for O'Connor, special and ontologically basic. It cannot be reformulated in terms of event-causation. However, O'Connor is committed to a physicalist metaphysics in which the exclusive constituents of reality are microphysical particles. O'Connor endorses the causal unity of nature, which holds that all of reality, even including the supposedly independent power of agent-causation, is a causal byproduct of event-causation at the microphysical level. The agent-cause is supposed to arise from the micro-physical level, in which there is only non-singularist causation, as an irreducible emergent property.[487] The upshot is a fatal tension. We are supposed to get a unique power of agent-causation, that is ontologically basic and able to exert a new form of causality capable of deliberate selection. At the

[486] Ibid., 72.
[487] Ibid., 110. O'Connor therefore endorses strong emergentism. See Pereboom (2001): 78.

same time, these unique capacities are completely reducible, and causal byproducts of, event-causation at the microphysical level.

The Significance of Character in a Metaphysics of Agency, and Its Lack in Contemporary Libertarianism

One concern for a metaphysics of agency is the need to appeal to an enduring character. Intuitively, if not factually, we sense that free action cannot be arbitrary—a mere whim of some indeterministic process. Action is something we do, not something that happens to us, which is why the same action can be the basis of different descriptions under which it is, and is not, an action.[488] Also, though, we sense intuitively that a free action does not arise through mechanistic causal process originating from outside the person. The essential complement is an enduring character of a person—a set of values, habits, virtues, and stable dispositions—in virtue of which a chance process, or a mechanistic process, involving beliefs and desires, becomes one's own.[489] Mere events, independent from a character, are incompatible with agential control.[490] An event in which an agent exerts control itself requires an account of how it is rooted in the agent,

[488] See Elisabeth Anscombe's example of all the different descriptions the action of moving one's arm up and down might take on. Elisabeth Anscombe (1957): 40.
[489] A comparable notion to *character* can be found in Frankfurt's notion of a well-integrated agent that is behind, i.e. has second-order desires that support, the desires on which she acts. Cited in Mele (2006): 170.
[490] Van Inwagen's consequence argument states that a deterministic event is the product of the past and laws of nature, which are not up to us, and so the event is not up to us. Cited in Mele (2006): 137.

and does not just happen to the agent. A character—Kant's term for it is *Gesinnung*[491]—at the root of processes involving the events of beliefs and desires, allows us to understand an individual choice in terms of agential control and not mere detached sequences of events, and ultimately draws the person out of a mere spectatorship on his or her own actions. A character secures imputability, since a mere sequence of events either happen by chance, through, as it were, a neural roulette wheel of competing motivations[492], has roots in a deterministic series outside the agent, or requires an infinite regress of exertions by the agent in order to secure control. A character also secures rationality, since mere events that happen to an agent, rootlessly and by chance, are only blind processes—i.e. they are not rationally guided.[493]

So, for example, it is one thing to claim that person *Y* stole the bike in virtue of a causal process involving the sight of the exposed bike, the stimulation of the desire for the bike, and the belief that one would get away with it (all psychological events). If we adopt the Davidsonian story of action, these causes of action, serving as a kind of practical syllogism, constitute its rationality.[494] But, rationality cannot be mere causal explanation, because the so-called agent, represented solely by the series of events, is indistinguishable from the impersonal causal processes happening around the agent. Action in this example is a merely reflexive response to external pressures, and the fact

[491] Henry Allison has an excellent discussion of *Gesinnung* in Henry Allison (1990): 136. An agent's character is the underlying set of maxims shaping individual exertions of *Willkür*.
[492] See the metaphor in Mele (2006): 114.
[493] We see the notion of character at work in Kant's notion of radical evil. Radical here refers to the notion of a root, instead of severity. Radical evil means that we have an ineliminable susceptibility to temptation that is the ultimate basis for all immoral actions. The incorporation of character means that individual immoral choices are not just sequences of events, but grounded in the nature of the agent to be susceptible to temptation. See Allison (2001): 230 for a discussion of radical evil.
[494] Davidson identifies reasons for actions in terms of causal explanations involving beliefs and desires in Donald Davidson (2001): 3-5.

that stealing the bike is mediated by drives internal to the agent's psychology makes no difference. Whether or not the agent steals the bike depends merely on quirks of the agent's psychology and environment over which the agent exercises no control. Using Kant's example of a malicious lie, the Davidsonian story focuses only causes having to do with the empirical character—bad upbringing, bad company, a temperament insensitive to shame, etc.—while neglecting the unconditioned causality of reason which makes the act imputable to the agent. Assuming only an event-causal metaphysics, the fact that one agent stole a bike, or told a malicious lie, and another does not in the same circumstances, is ultimately based on luck, i.e. the luck associated with the contingent factors of one's causal situation.[495] Luck is not rational, since the brute fact of one's causal situation lacks any normativity associated with rationality. Instead of choosing to steal because the agent considers it expedient, or even right, the Davidsonian agent steals it, in the final analysis, because that is the outcome of the causal factors coming together in a given situation. Any judgment of expediency, in a luck-driven scenario, cannot be the driver of a decision, but is instead controlled by non-rational circumstantial configurations. Alexander Baumgarten, whose metaphysics played an important part in Kant's philosophical career, defines a spontaneous action as one over which an agent controls the sufficient cause.[496] In a situation involving mere events, without an

[495] Luck objections to theories of freedom go back to Hume, who worried that any indeterminism incorporated into agency would sever a connection to the agent that would be the basis for imputability. See Derk Pereboom and Michael McKenna (2016): 237.

[496] Alexander Baumgarten (2013): Section 704. Kane similarly stipulates that a free agent must be in charge of the *arche* of her action in Kane (1996): 73.

underlying character, the agent cannot be the sufficient cause of action, since the action is a merely automatic byproduct of external and internal pressures.[497]

In sum, character plays a vital role in addressing what is known as the disappearing agent objection. An agent does not self-determine its own action if the action is the result of a sequence of psychological events, whether the sequence is deterministic or indeterministic. If the action just happens, as in an indeterministic process, or is the exclusive outcome of a deterministic process, then the agent has disappeared from the process of action.[498] Even if we identify an event involving an exertion of will leading to action, we need some way to articulate the mechanism by which the agent is in control of its will. A character provides an underlying basis in virtue of which the agent adjudicates its various psychological drives, and so settles its action, instead of disappearing. A disappeared agent represents the problem of the indifference of equilibrium, in which determining grounds for a choice have been removed and an agent has no preference biasing it in favor of any particular alternative. The action, detached from the agent's character, then occurs by chance.[499]

*Event-causal libertarianism and the presence of character***.** If action involves only probabilistic causation that allows the agent equal chance of going in alternative directions, then we are in a state of indifference of equilibrium that is not rooted in an agent's character. ECL identifies action as that which is caused psychologically, as a

[497] In the contemporary literature, the objection that certain theories of free will do not allow for an agent to settle an action, and instead delegate this ability to automatic processes, is called the disappearing agent objection. See Derk Pereboom (2014): Chapter 2.
[498] The disappearance of the agent occurs, I think, regardless of where one locates indeterminism—i.e. either in a conflict of the will itself, as in Kane, or in prior non-actional psychological processes leading up to the action, as in Chris Franklin. Franklin (2011): 205.
[499] Kant addresses the problems involved in indifference of equilibrium in *ND* 1:402.

result of desires and beliefs of which the agent is conscious. Even though a conflicted decision is resolved through an indeterministic process, ECL theorists will claim the final decision is at least the product of conscious choice.[500] The phenomenon of deviant causal chains exposes the insufficiency of specifically mental causation in differentiating action from non-action. What might happen is that a certain belief and desire combination causes some sort of behavior, and so we meet all the standards of action established in ECL, but the behavior is clearly not an action. For example, an agent desires his uncle's fortune, and believes he can get it by driving to the uncle's house and killing him. The agent runs over a pedestrian, who happens to be his uncle. What happens in this macabre example from Chisholm shows the insufficiency of the event-causal account: though the relevant beliefs and desires causes this individual to kill his uncle, this killing was clearly not an action but an accident.[501]

A subject, *S*, causes an event, *E*, in virtue of another event, *C*, in ECL. There cannot be direct causation between the subject and the action in ECL, because it rejects substance-causation. But, this mediating event, *C*, has to be a special sort of event if *E* is really to qualify as action. If I knock over a lamp because someone pushed me, knocking over the lamp was not my action. Event *C* has to be an action of mine if *E* is ultimately to be my action. But, if action consists in a kind of causation, as ECL holds, then we have an infinite regress problem. For *C* to be an action, it has to be caused by the events related

[500] So, Balaguer claims that we have authorship and control over an action even though it is the product of an indeterministic process. Mark Balaguer (2010): 78.
[501] The example comes from Roderick Chisholm (1966): 11-44.

to the agent, and these agent-related events must also be actions, and so caused psychologically, etc.[502]

The problems of deviant causal chains and an infinite regress shows the tendency of event-causal accounts of action to blur the distinction between mere accidental causal sequences, playing out in virtue of the quirks of an impersonal causal sequence, and deliberate action. The problem with accounting for action in terms of events is that it fails to secure imputability and rationality, regardless (perhaps especially!) if we introduce indeterminism. Beliefs and desires may cause an action, but even if the agent ends up killing his uncle knowingly and directly in the home, we are still dealing with an indeterministic series over which the agent has no control, and which are therefore not imputable.[503] Nor are the actions in ECL rational: the fact that an agent happens to have certain desires and beliefs that, through probabilistic causation, issue in a certain action, is a mechanical process detached from rationality. We need a richer metaphysics to lift us out of a merely mechanistic process, in which effects flow from their causes without any sort of reciprocal effect on their causes, to a teleological process specifically characteristic of intentional action. Intentional action involves not only the production of an effect (e.g. practical reason sets out to go out to lunch), but also an interplay of an effect on the cause (e.g. the idea of going out to lunch attracts the agent's rationality, as a prudent or moral thing to do).[504] An unfiltered transition from mental events to action is the kind of process an animal could undergo. Reason provides a depth, in which the flow

[502] Carl Ginet (1990): 7.

[503] Pereboom agrees: indeterminism means that the agent cannot be the source of the action. Pereboom (2001): xvi. ECL meets epistemic conditions for freedom—the agent is conscious—but not a control condition. Timpe (2013): 17

[504] As Kant puts it, the desired effect acts as cause in virtue of its concept, which serves as an end for the agent, and so the ground of possibility for the effect. *KdU* 5:220.

of events not only flows from cause to effect, but also involves a reflection of the effects back on the cause in light of certain normative standards.[505]

Character can provide the depth needed for the teleology of action. Character is the site of a set of values—maxims, as Kant puts it—in virtue of which an agent can intervene within the sequence of psychological events, evaluating these sequences in a way that allows the agent to truly settle his or her action. The metaphysics of ECL are Humean in its rejection of a substantial character underlying events. For ECL, the question of whether or not we have free will is an empirical question about whether or not there is the appropriate level of indeterminism leading up to the moment of choice.[506] Balaguer, who is a prominent defender of ECL, defends the sweeping position that any metaphysical questions that are not amenable to empirical verification or logical analysis are just empty questions. If there is mention of character in ECL, as there is in Kane, it can only be a token mention. Efforts of will arise from character for Kane—i.e. competing sets of reasons arise from who we are—but the indeterminacy of efforts of will means that there is an indifference of equilibrium in which the action ultimately issues by chance.[507]

ECL subscribes to the principle, in the interests of an economical empiricism, that the metaphysics of agency in a libertarian theory should not include any mechanism beyond what is needed in compatibilist accounts.[508] But, this metaphysics, which fails to incorporate an underlying character, is perfectly compatible with a merely instrumental

[505] See J. David Velleman (1992): 461.
[506] Mark Balaguer (2010): 22-4.
[507] Kane compares these indeterminate efforts of will to a quantum particle that may or may not penetrate a barrier. Robert Kane (1996):128
[508] Kane calls this principle the free agency principle. Robert Kane (1996): 117.

kind of reason that only finds ways to achieve the ends set for it by passions. The metaphysics of ECL cannot accommodate the power of reason to determine itself. Since the will in ECL is indeterminate[509]—like a quantum particle that may or may not pass through a barrier—it can only passively respond to the outcome of the chance process.

The key distinction at play here has already been made in Kant's *Religion*. The disposition to *personality*, and so responsibility, differs from the disposition to *humanity*, and rationality, in virtue of the fact that the mere faculty of reason as such does not include the unconditioned power of practical self-determination. Even a hyper-rational being might respond only to empirical incentives acting on it as external stimuli.

Personality, as opposed to mere humanity, consists in the possession of an unconditioned power of choice that is independent of any external or internal stimuli.[510] It is the unconditioned power of choice, in which an agent forms its own volitional determinations as opposed to passively receiving them, that is needed to rescue the ECL agent from illusory agency. The ECL agent has only an instrumental rationality that enables it to pursue whatever intention falls out from the indeterminism of the torn decision.

In the *Religion*, Kant distinguishes between two meanings of the word 'deed' (*That*): it can mean a supreme maxim that provides a formal structure for individual acts, or it can mean an individual deed occurring in concrete circumstances (*factum*

[509] Notice I am talking about indeterminacy of will, not indeterminacy of *situation*. The latter (e.g. radioactive material may not give cancer) *is* compatible with an unconditioned power of reason. Pereboom (2001): 43.
[510] *R* 6:27.

phenomenon).[511] ECL incorporates actions in their material aspect, i.e. in terms of the concrete empirical objects to which they are related, but ECL insists on a metaphysics that is hostile to the idea of a supreme maxim that oversees certain individual actions.

Non-causal libertarianism and character. We are not going to be able to ground action in character, in virtue of which competing psychological motivations are adjudicated in terms of an unconditioned power of choice that controls the sufficient reason of an action, by detaching the power of choice from any causal structure and making it a free-floating and intrinsically active capacity. Without some sort of causal relationship by which the agent can control his mental power to act, the mental power will behave with full detachment. An NCL agent, just like an ECL agent, is an unusual species of agent that would not be appropriately included in human participatory attitudes. Participatory attitudes in interpersonal interaction attribute actions to persons. We can provide a metaphysics appropriate to these attitudes by incorporating actions into a larger metaphysical structure that involves an intelligible character. Participatory attitudes presuppose that the actions of people around us are not free-floating events. We do not interact with a mental property that commits acts, but with the person in charge of their mental properties. There is a character to which individual actions are attached, in virtue of which individual actions (*Handlungen*) in empirical situations flow out from general rules contained in maxims that provide a fundamental orientation to one's will. But, NCL denies this structure, and can only attribute action to the person in a way that is empty of explanatory content, since it abjures any causal structure associated with action. NCL leads to an objective attitude instead of a participatory one, in which we view action

[511] *R* 6:31.

in terms of spontaneous firings of mental powers that are detached from agential control.[512]

NCL incorporates rationality—it is not as though mental powers act with utter arbitrariness. We can append purposiveness to an uncaused mental action in virtue of a teleological structure. NCL, since it accounts for action in a way that is detached from a causal structure, locates action in the mere taking of a propositional attitude. Propositional attitudes, for NCL, are choices that have an intrinsic activity. A propositional attitude might be oriented towards a future state or purpose. The teleological structure is not a causal structure, but nevertheless it saves the propositional attitude from complete arbitrariness, which seems incompatible with action.[513] Still, if there is no causal connection of the mental power to the agent, the addition of a teleological structure does nothing to alleviate the problem of agential control. The agent in this case would be passively suffering the occurrence of teleological structures that are uncaused and not grounded in the agent's character.

The mere presence of a representative state of a mental event with a so-called actish phenomenal quality indicates, for NCL, a capacity for practical spontaneity.[514] However, mere representative states are neutral with respect to the causal structure in which an agent is enmeshed. A being capable of representation could act in ways that are the inevitable product of causes completely out of its control. NCL is a libertarian theory and so it requires an indeterministic causal gap prior to the exercise of choice by the

[512] The definition of participatory and objective attitudes comes from Peter Strawson (1962): 1-25.
[513] Stewart Goetz (2008): 19-22.
[514] The importance of the phenomenal aspect in NCL makes it a mirror of epistemological TI, in which our freedom is based on taking ourselves *as if* we are deliberating over our inclinations in terms of maxims, and not free in an ontological sense. See Pereboom (2001): 41.

mental power, but this does not alleviate the problem of a disappeared agent in NCL. A being with what, merely phenomenally, are acts, could be subsumed in mechanistic causation or have volitional capacities equivalent to a roulette wheel. The inference, then, from phenomenal acts, regardless of causal structure, to genuine practical agency, is a defeasible inference in light of objective facts about underlying causal structures. A metaphysics of character can provide the causal structure that involves the agent in determining its own actions, rather than falsely accepting as actions mental states that have an independent causality and are only actions in virtue of a qualitative feel.

*O'Connor's naturalistic agent-causation and character***.** It is the judgment of this dissertation that agent-causal libertarianism, in which the agent directly causes its action without mediation by sequences of events, best incorporates an agent's character and so avoids the problem of a blind and detached sequence of events, from which the agent has disappeared, in both ECL and NCL. However, the metaphysics of agent-causation are tricky: how, for instance, is an agent supposed to directly cause an action, in a way that both has explanatory content and is irreducible to causation by events?[515] O'Connor is committed to a naturalistic metaphysics in which all of reality can be accounted for in terms of the potentialities inherent in the different relationships and configurations at the microphysical level. At the same time, O'Connor promotes the idea that an agent-causal power, capable of self-determination independent from events, is able to arise from a microphysical level which lacks this active power.[516]

[515] Agent-causal accounts are sometimes accused of lacking explanatory content. See Derk Pereboom (2014): 54.

[516] O'Connor claims that, though reality is causally united with the microphysical level, it is not constituted by this level. Timothy O'Connor (2000): 108.

O'Connor's view is equivalent to a reversal of a Kantian metaphysics in which an unconditioned and spontaneous power of choice cannot exist in the empirical time-series. There can be no empirically unconditioned in the time-series, because the fact of temporal succession requires a causal law that grounds an event in an asymmetric relationship with a prior event, which in turn accounts for its objective place in the time-series.[517] For the Kantian, the empirical series is not supposed to be ontologically fundamental: instead, it is supposed to be indeterminate, metaphysically derivate from an absolute reality that is detached from human sensibility and thought through the pure categories. The empirical series is subject to a merely contingent form of human sensibility that is not an absolute determinant of reality in itself. If an unconditioned power of causality is supposed to arise in an ontologically fundamental empirical level, then what is uncaused—the unconditioned power—must paradoxically also be caused, since it is only a causal law that can situate something in objective temporal succession.[518]

Though laden with Kantian jargon, the paradox of an unconditioned arising from a conditioned series is relevant to O'Connor's notion of emergent properties. An agent-cause for O'Connor is supposed to exercise an ontologically basic and irreducible mode of causation, that consists in an active power of self-determination, that nevertheless can be accounted for completely by a mode of causation at the microphysical level that is only passive and mechanistic.[519] This is a kind of magical emergence—an unconditioned

[517] Kant claims that there can be no empirical unconditioned at *KdrV* A531/B559.
[518] This is the argument of Kant's Second Analogy. *KdrV* B233.
[519] O'Connor notes the objection that the agent-causal power, because of his causal unity of nature thesis, can be accounted for entirely in terms of the causal potentialities of the microphysical level, and so the emergent property is really illusory. Timothy O'Connor (2000): 112.

causality that is also conditioned—that involves a new and irreducible power that somehow arises from, even though it does not track the causal powers of, the microphysical level from which it arises. It makes sense that macro-level phenomena derive their power from the causal potentialities of the microphysical level, but O'Connor wants to both acknowledge the possibility of an ontologically irreducible new agent-causal power while still trying to maintain naturalistic credentials involving a commitment to causal unity with the microphysical level. [520]

The metaphysics of character cannot arise from the empirical series in time, then, but must instead underly it. If we introduce an underlying character, we can add a dimension to the event-causal sequence in virtue of which the agent adjudicates the flow of psychological events from external stimuli, taking a second-order evaluative stance on beliefs and desires.[521] Agents respond differently, then, to contextual forces, not because of accidental and mechanical processes, but in virtue of rational judgment. However, the character underlying individual actions creates its own problem for a theory of agency. If the character is fixed, then the character becomes just another accidental factor in the automatic causal sequence leading to action. Whether or not one steals the bike is just a matter of luck—not rationality—having to do with one's inborn character. If we claim that one chooses one's character, we seem to attribute to the individual agent a fantastic power to determine its entire history of individual acts from a God's-eye perspective, in

[520] The idea that higher-level causal explanations must track causal powers at the microphysical level is Jaegwon Kim's causal inheritance principle. Jaegwon Kim (1993): 309-35.

[521] Kant gives an account of character in his *Religion*, where he claims that the character—or, disposition—is the subjective ground of the adoption of maxims. This is not a foundational choice in time—some sort of myth of Er scenario in which the individual selects his whole lot in some mysterious original position in the afterlife. Instead, the choice of the subjective ground of all maxims has its origins in reason outside of time. *R* 6:25.

which the individual selects the underlying nature at the root of all its acts over the course of her life.[522] Also, the idea of a character seems to simply reintroduce the original problem of agential control at another theoretical level. If there is an event in which the agent chooses its character, in virtue of what does the agent control this event? If the event is determined or the product of indeterminism, it is not subject to rational control. The problems of character then force one back into trying to articulate freedom in terms of isolated exertions of the will, detached from an enduring character, and then the conundrum of imputability and rationality returns in its old form.

I will now show how the competing concerns of both stability, which allows us to connect individual choices to a character and so install agential control, and freedom, in virtue of which an agent is not controlled by its character and is able to reinvent itself, are best handled by the metaphysics of transcendental idealism.

Metaphysical TI As a Framework for Character

Metaphysical TI is a powerful paradigm which can fuse an underlying character to an event-causal series, without at the same time making the character into some fixed nature outside of agential control or into a weird foundational choice in time of all one's future maxims that itself requires a regress of higher characters to account for its own imputable origination. In metaphysical TI, the sequence of events in the Davidsonian

[522] Henry Allison compares this foundational, God's-eye choice of the ground of all one's actions to Plato's myth of Er from the *Republic*, in which souls in the afterlife are allowed to select their lots prior to actually living them. Henry Allison (1990): 139.

practical syllogism of beliefs and desires leading up to action is not an ungrounded absolute. It is in a time-series that is indeterminate, because it depends on a relationship to a contingently structured sensibility, and on an intelligible order that is independent of this contingent sensibility. While the empirical psychological drives in time may be due to the contingency of our circumstances, their indeterminacy within the framework of transcendental idealism means that luck is not the controlling factor. Instead, the indeterminacy of appearances is complemented by the determinacy of the intelligible level. The intelligible order can serve as the site of maxims, that account, as a kind of substantial form[523], for the structure of causation at the empirical level. Two individuals may both covet an unattended bike they see, but differ in their responses, not because of mere luck having to do with the quirks of their inner and outer contextual situation, but because of maxims that structure their character and which to varying degrees privilege or suppress different psychological inclinations.[524] As Kant notes, the intelligible character is a transcendental cause of the empirical character—the transcendental status is importantly different from a preceding cause in time—so that the empirical character is the "sensible sign" of the intelligible character.[525] Moreover, the fact that the intelligible character is a transcendental cause outside of time allows us to avoid the theoretical difficulties of a fixed character that just reinforces the agent's subjection to fortuitous

[523] The idea of a substantial form is an Aristotelian idea that was very important to Leibniz, and I think it can be usefully applied to understanding the way the intelligible character interacts with the empirical character. In Leibniz's middle years at least, he thought that a body could not be accounted for solely in terms of extension, and that it needed an intelligible structure to account for its unity and the form of matter that composes a body. For an account of substantial form in Leibniz, see Robert M. Adams (1994): 308.

[524] Maxims, which, as intelligible realities, do not function as antecedent causes in time, serve as a structuring causes. Structuring causes account for the linkages between different events. For instance, pressing the doorbell leads to a ring because of the structuring causes represented by the wiring that allows electricity to flow so the doorbell can work. See Timothy O'Connor (2000): 53.

[525] *KdrV* A546/B574.

circumstances. An intelligible cause outside the time series allows us to exempt it from the causal law that structures objective succession in Kant.[526] A Kantian agent is able to reinvent the structure of its maxims according to a rational judgment, and not in terms of antecedent cause, as if the character were some natural effect arising in time.[527] The intelligible structure of the agent represents an unschematized form of causality that is a pure activity not bound to the series in time available to our sensibility.[528] Since the intelligible character is not some antecedent cause in time, we do not have to articulate it as a foundational first choice in some weird state of pre-existence, or to get into an infinite regress by trying to account for how we are to establish control over our own characters in repeated antecedent events. It is, in fact, the unschematized timelessness of the intelligible character that allows us to combine the stability and continuity of a character to establish agential control (a kind of Aristotelian *hexis*, i.e. deep dispositions of character) with a Sartrean *projet fundamental* (some original free choice that unites the various empirical behaviors of the agent).[529]

Of course, such a metaphysical picture, that combines an indeterminate series in time with a metaphysically fundamental intelligible structure that lifts the series of events in time out of a merely impersonal contingency, is unavailable to the transcendental realist. If one is a transcendental realist, and views space-time as outside of one in a

[526] I am referring to the argument of Kant's Second Analogy, in which he claims that we need a causal law to validate succession as truly objective, and not just as based on the order in which we perceive things. *KdrV* B233.
[527] Kant distinguishes between the mere effect at the intelligible level, which is just an unschematized activity, and an occurrence in time, in *R* 6:39.
[528] The schemata are supposed to unite the heterogenous elements of the pure and general concepts with particular sensible intuitions. See *KdrV* A138/B177.
[529] The tension between the stability of character and the freedom of self-invention is discussed in Henry Allison (1990): 139.

transcendental sense—i.e. as existing in an absolute and mind-independent way, and not merely as appearing as outside of one—then this fusion of the empirical series of desires and beliefs, and an intelligible character, is utterly impossible. If events in space-time are metaphysically fundamental, then there can be no non-spatiotemporal intelligible ground of these events.

Contemporary libertarianism embraces a transcendental realism about spatiotemporal events. This transcendental realism views space and time as pertaining to the structure of reality independently of intuition—i.e. space and time have *absolute Realität*.[530] What is at issue between contemporary libertarianism, and the sophisticated Kantian theory of free will I have developed, are competing metaphysical paradigms that differ with respect to what they take to be most fundamental. Since contemporary libertarianism, in various ways, tries to articulate a metaphysics of agency in terms of the spatiotemporal series of events, it is clear that, for it, what is most fundamental in metaphysics is the spatiotemporal series. Contemporary libertarianism will even use a naturalist metaphysics as a test of a metaphysics of agency: our metaphysics of agency is to be judged in terms of how well it conforms to the nature of the world as portrayed by cutting edge natural science. Derk Pereboom rejects agent-causal theories because they

[530] It is important to recognize that transcendental realism is an umbrella term that applies to different metaphysics of space and time. Both the Leibnizian and the Newtonian views of space and time are examples of transcendental realism, even though the former views as inhering in thing and the latter views them as absolute containers of things. The commonality is that both Newton and Leibniz attribute to space and time a reality that is independent of intuition. *KdrV* A36/B52.

entail interruptions of deterministic laws at the physical level, but there is no evidence for such disruption.[531]

The result of the presupposition that events in space-time are fundamental in contemporary libertarianism is that there is only one kind of grounding relationship: an antecedent event in time. But, a Kantian metaphysics views these events in time as schematized versions of otherwise pure categories, applied to the spatiotemporal structure of a contingent form of sensibility that is not an absolute measure of the nature of reality. In his distinction between origination according to reason and origination according to time in the *Religion*, Kant shows that causality has a pure meaning as a power producing an effect, which does not necessarily have to do with some causal sequence of events in time.[532] Contemporary libertarianism, rather than accepting that sensible representation of events in space-time involves a contingent relationship to a particularly human kind of sensibility, instead accepts a Humean supervenience in which what is most ontologically fundamental, and determinative for the configuration of the macro-structure, are particulars at points in space-time.[533]

The diversification of grounding relationships in Kantian metaphysics can solve a host of problems. For instance, if we try to insert an agent-cause in an exclusively event-causal empirical series, we end up with another luck problem in explaining why the agent in one scenario, and not another, performed a given action, even given the same causal series leading up to the action. An agent-cause is supposed to be an active self-

[531] Derk Pereboom (2014): 66. Mele also argues there is no evidence for agent-causation in Mele (2006): 134. I discuss how we can integrate agent-causation into deterministic empirical lawfulness in Chapter 5 of this dissertation.
[532] *R* 6:40.
[533] David Lewis, (1994): 473-90.

determining cause, and so able to act differently even given an identical causal series. But, if all we have as causal mechanisms are events in time, there is no way to account for different outcomes, given an identical event-causal series, except in terms of luck. One might point to an agent-cause to account for different outcomes, that is somehow independent of the event-causal series, but if one's metaphysics permits only the exercise of causality through events in time, how is this agent-cause supposed to operate other than through an event in time?[534] If we try to locate agent-causation in a particular event—e.g. the event of the agent causing event *e* at time *t*—and attempt to seal off this event from determination by previous events by introducing indeterminism, then we again have a luck problem in which the agent-causal event occurs by chance.[535] If the fundamental entities in our metaphysics are particulars occupying points in space-time, the invocation of an agent-cause will be empty.[536] We can assert that an agent-cause generates an action, independently of a causal series, but without some articulation of substance-causation without events, it is not clear how the productive power of the agent is supposed to work. A Humean metaphysics could only account for this productive power in terms of another event, which itself would require some connection to the agent, or no event at all, which would re-introduce the luck objection.

The metaphysics of TI and ECL. In my analysis of ECL's accommodation of character, I mentioned how ECL proposes a sequence of action leading from the *S*, the subject, to the action, *E*, through a mediating event, *C.* If the mediating event, *C*, is supposed to make *E* an action (since *C* is appropriately psychological), then it is fair to

[534] See the discussion of the luck objection against agent-causation in Derk Pereboom (2014): 52.
[535] Carl Ginet makes this objection, cited in Ibid., 54.
[536] Alfred Mele makes the point about the emptiness of the agent-causal explanation, cited in Ibid., 54.

ask in virtue of what is *C* an action? *C* could be a non-actional external prompt, such as someone pushing one. Introducing another psychological cause of *C,* to account for what makes it an action, only postpones the problem, as we have to account for how this new event is not also an external prompt. The proper resolution of this problem is a new kind of ground, different from an antecedent event in the empirical series. Here, there is a causal relationship between events related to the subject—*C*—and an event—*E*. What makes *C* an action is not another layer of agent-related antecedent events that connect *C* to the agent, instead of an external force, like a push. What makes *C* an action is the fact that the *C*-*E* connection occurs in virtue of a maxim acting as an intelligible ground and a second-order principle that governs the linkage of agent-related events and actions in a multitude of individual acts. The intelligible ground, since it is not in the time-series in the empirical domain of schematized categories, can be accounted for in terms of a rational choice and not in terms of another antecedent event.

One objection to accounting for action in terms of an agent-cause is that, in bypassing events and opting for direct causation by an agent qua substance, an agent-cause cannot account for the specific time-determination of its action.[537] The solution I am offering here, though, does not omit events in favor of direct causation by an agent. The *C*-*E* linkage of events is still intact. What my account adds is a grounding of *C*-*E* in an intelligible structure, thought through the intellectual synthesis and detached from human sensibility, that accounts for why certain event linkages occur in certain agents

[537] We find this objection in Eric Watkins (2005): 243.

and not others.[538] ECL, though it can account for the temporal determination of an action through an equally temporally determinate cause, has a difficult problem in its assumption that the appearances to a cognitively limited human subject are determinate and metaphysically fundamental. This is a form of dogmatic empiricism that itself is immodest and even contradictory, since there is no empirical basis for the belief that empirical content is the exclusive content of reality.[539] This dissertation adopts an immaterialist metaphysics in which the empirical domain available to human cognition is a merely subjective-relative domain that can be complemented by a structure of reality in itself that transcends the limits of human sensibility.[540] While schematized cognitions of human sensibility are in space and time, schematization under spatiotemporality does not exhaust the ontological structure of reality, as there are non-empirical aspects of the self that complement the temporally determined event-causal order. The non-empirical domain of the self does not have to replace the event-causal order, in a way that subverts its temporal determinacy—i.e. introduce a substance-cause that produces events without an event. The substance-cause can instead serve as a ground because the empirical event-causal order is metaphysically relativized to human sensibility.

Critics of agent-causation also claim that a direct relationship between agent qua substances and actions eliminates the possibility of reasons for action. If reasons consist in a causal sequence of events leading up to action, as ECL claims, then one can only

[538] My account is therefore similar to that of Randolph Clarke, whose version of agent-causation presents the agent not as directly causing an action, but instead causing the fact that the action is done for certain reasons. Randolph Clarke (1993): 191-203.
[539] See Kant's criticism of dogmatic empiricism at *KdrV* A471/B499.
[540] See the different levels of being Karl Ameriks identifies in virtue of transcendental idealism in Karl Ameriks (2012): 243. These different levels are empirical features, nonsensory features of empirical beings (e.g. free will), relational feature of non-empirical beings (e.g. God), and the intrinsic nature of transcendent being.

conclude that direct causation of an action by a substance, which does not consist in events, is irrational. However, in the multi-dimensional agent-causal account in this dissertation, the antecedent grounds in the empirical series remain, and so desires and beliefs, arising as events in time that cause an action, can still serve as reasons, though they are not sufficient causes. The intelligible ground provides a higher-order metaphysical structure in virtue of which certain reasons are connected with certain actions, instead of immediately manifesting themselves in action. The higher-order character is not an antecedent ground in the causal series—i.e. the character, existing at a certain time, exerts some kind of force so that later belief-desire events ultimately end up in a certain action—but instead a higher-order *logical* ground that structures causal relationships. For example, one's maxim that one should always perform household chores serves as a logical structure in virtue of which the event of a tempting desire to scroll through the internet does not lead to the neglect of one's duties.

ECL, though, rejects an immaterialist metaphysics, that relativizes the domain of mental events to human sensibility, and allows for the agent-cause as a non-empirical power. ECL's rejection of immaterialist metaphysics comes from an empiricism that rejects any kind of non-empirical causation as obscure, since the only source of cognitive content is empirical impressions. Non-empirical powers have brought down upon libertarian theories the charge of being unintelligible, and unable to compete with the developing picture of the self coming from the sciences.[541] The cost of this aversion to

[541] Robert Kane adopts an empiricist principle of parsimony he calls the Free Agency Principle, in virtue of which we introduce metaphysics into our account of agency only if it is required by both libertarian and non-libertarian theories. Robert Kane (1996): 121.

metaphysics, though, is the disappearance of an agent as sequences of events involving indeterminate efforts of will take over the agent's behavior.

*The metaphysics of TI and the naturalistic metaphysics of O'Connor***.** O'Connor defends a form of agent-causation in which the agent has a special power of self-determination. Whereas event-causation involves a merely passive transfer from circumstances to effects—a transfer of activity based on the nature of something, such as occurs when water, in virtue of its mass and the force of gravity, passes downstream—agent-causation is an irreducible and special genus of causation that is able to freely self-determine itself to choose among different possibilities.[542] While promoting this special genus of causality that is not further analyzable in terms of events, O'Connor seeks to preserve a commitment to naturalism. Concerned to rebut the suspicion that agent-causation is not consistent with the emerging scientific picture of the world, O'Connor promotes the thesis that all macro-physical phenomena are causally united with naturalistic micro-physical phenomena.[543] Agent-causation, then, even though it occupies an irreducible causal genus that exercises an unconditioned causality to choose between various alternatives, nevertheless arises from the event-causal order at the microphysical level. O'Connor introduces the idea of an emergent property in order to reconcile the seeming contradiction between unconditioned self-determination arising from passive event-causation. When the event-causation at the microphysical level reaches a certain level of complexity, the theory goes, the laws that characterize it undergo a discontinuity

[542] O'Connor (2000): 72.
[543] He calls this the causal unity of nature thesis. Ibid., xv.

as a new property of agent-causation emerges, able to go against the laws typical of the base properties in the event-causal order.[544]

O'Connor spends a lot of time defending his idea of emergent properties against objections[545], but it is the judgment of this dissertation that the idea that an irreducible genus of unconditioned causality is also causally united with event-causation carries a heavy burden of proof. It seems impossible that an "ontologically basic"[546] species of causation—ontologically basic in the sense that agent-causation is a genuinely new form of causation exercising causality in a top-down fashion—could also arise from causation at the microphysical level. The problem of explanatory exclusion dismisses higher-order and emergent causal powers in light of the fact that first-order causal powers at the microphysical level sufficiently account for all macro-level phenomena, and therefore higher-order properties are causally idle. If nature is causally united, and fully based on the potentialities of the microphysical level, explanatory exclusion entails that there could be no emergent form of causation with genuine causal efficacy (i.e. it could only be an epiphenomenal perspective). O'Connor's emergent properties seem to be magical emergent properties, which exercise causal power without at the same time having any basis in the micro-physical levels which constitute them.[547]

In Kantian terminology, O'Connor's agent-causal power is an unconditioned power that arises from the empirical series of conditioned events and their conditions unfolding over time. The transcendental realist views that time-series in a dogmatic

[544] Ibid., 113.
[545] For instance, in this paper: Timothy O'Connor (1994): 91-104.
[546] O'Connor uses this term at O'Connor (2000): 111.
[547] For a discussion of these problems, see Derk Pereboom (2001): 75 and Jaegwon Kim (1993): 326.

fashion, as pertaining to the thought of things in general and not merely to objects of sensibility. Accordingly, as reason thinks the unconditioned, the transcendental realist must view the unconditioned as tracking the temporal series. O'Connor's embrace of the causal unity of nature—the idea that all macro-physical phenomena can be accounted for in terms of causation at the microphysical level—can be understood as a form of transcendental realism, insofar as it posits the spatiotemporal series as the most fundamental level constraining things in general as an ontological condition. But, reason's aspiration towards the unconditioned, when forced to track the temporal series, runs it contradictions. The unconditioned in the time-series lacks a condition to account for its presence in the time-series, and so it disrupts the unity of experience.[548] A location in the time-series—i.e. the fact that the unconditioned occurred at one time, and not another—is itself a contingent feature for which reason naturally requires a condition.[549]

The person, as Kant noted in his final work, *Opus Postumum*, connects the two worlds, i.e. the empirical and the intelligible world.[550] In O'Connor, there is just one world from which the person is supposed to emerge, but in Kant, the deepest faculty of the person—i.e. her power of transcendental freedom—touches on the intelligible domain outside of the time-series, while at the same time the person manifests herself on the empirical level as partaking of the series of antecedent grounds in time. Since O'Connor's idea of emergent properties carries such a heavy burden of proof, it would be better to reverse the determinate/indeterminate relationship. That is, instead of making

[548] We see the problem of the violation of the unity of experience in the antithesis argument of the third Antinomy. *KdrV* A445/B473.
[549] *KdrV* A460/B488.
[550] Cited in Dieter Henrich (2003): 52.

the microphysical event-causal level fundamental and determinate, insofar as its lawfulness does not require a certain kind of complexity to characterize it (unlike the indeterminate agent-causal power that only arises as an emergent property given a certain level of complex organization), we can make the microphysical level indeterminate, insofar as it pertains to a relationship to a kind of sensibility with a specific kind of structure. If the microphysical level is indeterminate, its genus of event-causality does not have to constrain all kinds of causality that might characterize our ontology. We can instead posit the unconditioned at the intelligible level, in a determinate order that is detached from any particular kind of sensibility. In the Kantian person, then, the unconditioned cause is the ground of the empirical series, whereas in O'Connor's agent, the empirical series of conditioned causality is supposed to be the condition, paradoxically, of an unconditioned form of causality.

World-Constitution in Empiricist Libertarian Theories and Kantian Metaphysics

Pre-Kantian debates on the constitution of a world had to do with whether ideal—i.e. in thought alone—or real connections between things were necessary to establish different things as part of the same world.[551] Kant, in his *Inaugural Dissertation*, clearly sides with real connection as a necessary condition of worldhood, claiming even that it is the potential for real causal interaction, not actualized interaction, that is the necessary

[551] The opponent here is Leibniz's idea of pre-established harmony, in which participation in the same world of both body and soul occurs through a mere correspondence in virtue of eternal essences arranged by God, and not through any real causal interaction. See the discussion of these debates about worldhood in Eric Watkins (2005): 85-9.

precondition. Focusing on the form of potential interaction as fundamental to the world is part of the relegation we see in Kant of events in space-time, embodied in actualized causal relationships, to a metaphysically derivative status.[552] Kant goes on to claim that space and time are only the form of the sensible world, uniting only appearances, and so are less metaphysically fundamental than the subject-independent principle that unites the world in itself, independently of the way it appears.[553] Kant's position on the constitution of the world stood in opposition to that of Leibniz, who thought worldhood can be formulated in terms of causally isolated substances that do not interact with one another, but merely correspond with one another in harmonious fashion according to a divine coordination. Leibnizian relationships between things are merely ideal, and not real.[554]

At first debates about the constitution of the world may seem esoteric and unrelated to a metaphysics of agency, but in fact, since agency is part of the world, the larger metaphysics of worldhood have important implications for the metaphysics of agency. For instance, the idea of the unity of worldhood founded on a potential for interaction between substances, that is ontologically prior to a manifestation in the form of events, as well as a relegation of the spatiotemporal framework of events to the phenomenal world of sensibility, permits the sort of hybrid agency that spans both the intelligible and sensible worlds I have been developing. This hybrid empirical-intelligible agent not only is able to exercise an unconditioned causality that is not constrained by a causality in the spatiotemporal domain, but the hybrid agent is also able to exercise its

[552] See *Diss.* 2:390. Kant repeats the claim that a dynamical community is the necessary basis of empirical cognition of a spatial community in *KdrV* A213/B260.
[553] Ibid., 2:398.
[554] See the discussion in Watkins (2005): 25, 76.

activity in a way that does not have to manifest itself as an event. Just as the unity of merely potential interaction exercises an activity that is not yet an event, the intelligible aspect of the two-worlds agent is able to exercise an activity over its own empirical manifestations that is not itself an event. This new dimension of agent-causation through an activity, not localized in the form of a specific event, permits an avoidance of certain grave problems in event-based models of agent-causation. If we deal with a metaphysics in which the spatiotemporal framework of localized events are fundamental, and agent-causation must exercise itself in the form of an event, we face the dilemma of agent-causation through chance, in the form of an event that itself lacks a prior event, or an infinite regress of events, in which we continually appeal to other events to establish agential control.[555]

The form of possible influence that is the basis of worldhood for Kant is immutable, and so independent of the changing determinations of events in space-time.[556] This abstract and general framework, of potentially united substances that are ontologically prior to any manifestations in the form of spatiotemporal events, can receive specific application in the form of an integrated character I discussed in section two of this chapter. The character itself is not an event, but a stable set of dispositions, values, traits, etc., that endure over time. If we try to integrate the character in the form of a specific event, we simply postpone the problems of imputability and rationality the character was originally meant to resolve. In virtue of what, we have to ask, does the

[555] Watkins mentions this problem for the agent-causal models of O'Connor and Taylor at Watkins (2005): 19.
[556] Ibid., 2:390.

newly introduced event, of the impact of character, belong to the character of the individual?

Severing the potential interaction of substances in the world, in the manner of Leibniz, entails for the metaphysics of agency a causally idle will, which exercises no genuine control over the empirical manifestation of an individual's actions. The empirical manifestation of the Leibnizian self is a mere reflection that happens, through divine intervention, to correspond to the internal changes of the grounding monad.[557] The Humean worldview also severs simple impressions from any kind of metaphysical substructure, and would reject the metaphysics of the pre-Critical Kant (as well as the Third Analogy of the Critical Kant[558]) in which the form of potential metaphysical interaction is ontologically prior to the events in space-time. The Humean worldview undercuts ontological structures in the name of an empiricist parsimony that forces a grounding of all one's ontology in individual impressions, divorced from underlying substances or connective powers.[559] Taking simple events as metaphysically fundamental raises serious issues for how individual events are to participate in the same world, since simple events cannot even be the ground of continuous powers by which one event is united to another. This fragmented Humean world, of individual events popping into and out of existence separately from metaphysical powers or underlying substances in mutual interaction, has serious entailments also for one's metaphysics of agency, which I will now describe.

---

[557] See the discussion of Leibnizian metaphysics in Eric Watkins (2005): 25.
[558] In the Third Analogy, Kant argues that interaction between substances is necessary for the objective simultaneity of events. *KdrV* A211/B256.
[559] For insight into Humean metaphysics, specifically its rejection of substances, see David Hume (1949): 220.

If particulars at spatiotemporal points are ontologically fundamental, then we cannot articulate causality in terms of an actual unifying and metaphysical power, but instead in terms of mere associations of events that are psychological in nature—i.e. Humean customs that relegate causation to a psychological projections rather than to a metaphysical power.[560] A disunified world, in which a reductionism to particulars in space-time eliminates a causally interactive form that is not actualized in space-time, also cannot ground a character in our metaphysics of agency. A character requires a continuous presence underlying events spread out in the time-series, as well as a causal relationship that is not reducible to an event, in virtue of which certain events belong—i.e. are grounded in—to a certain enduring character. None of this can be grounded in a Humean metaphysics which privileges particular events, and so cannot include causal relationships involving grounds lacking the temporal determinacy of events. A Humean metaphysics is like a canvas on which one can splash drops of paint in any pattern of colors one likes, because there is no underlying causal structure that constrains possible patterns.[561] A metaphysics of particulars in space-time that lacks causal constraint cannot be the ground of anyone *doing* something: the mere occurrence of events, absent some special grounding relationships, merely happen in ways that are driven by luck.

A disunified world, which cannot even, by Kantian standards, constitute a world, because of its causal isolation, includes only events that do not constrain one another, because there is no underlying causal structure of the world. Such a false world cannot even account for the fact that certain causal powers are inherent to the identity of things,

[560] See David Hume (1999): 7.2.29.
[561] The metaphor of a canvas that accepts any pattern of color without any underlying structure comes from Gregg Rosenberg (2004): 143.

which entails that we cannot arbitrarily vary effects (e.g., certain stable chemical properties). A Humean world without causal connection can splinter into an infinitude of parallel universes, in which there is not a temporal unity in virtue of which individual events participate in the same temporal flow. Kant notes in his argument for the First Analogy that events can only arise in relation to a persisting substance, because mere empty time cannot provide a standard of orientation against which to assess temporal determination. Since persisting substances are the ground of the unity of all perceptions in a single temporal structure, removing them, and replacing them with events that also arise at a certain time, would fragment time into parallel tracks.[562] The succession of impressions alone, without an underlying substance, cannot provide a basis for unified time.[563]

The disunity of the world of Humean simple impressions ultimately carries over to the disunity of the agent, and I think a Humean metaphysics is actually at the root of the problem of the kind of stunted pseudo-agency we get in both NCL and ECL. The agents in both ECL and NCL are fragmented: the agent is a mere spectator as indeterminate efforts of will resolve themselves according to chance, or as intrinsic powers fire off without any sort of causal structure connecting them to the agent.[564] For the Kantian, the appearance of particular facts in space-time presupposes metaphysical interaction of substances as a more fundamental reality.[565] Accordingly, the Kantian can

[562] The discussion of a possible splintering of time into parallel tracks because we replace substances with changing events occurring in *KdrV* A189/B232.
[563] "For in the mere succession alone existence is forever disappearing and recommencing, and never has the least magnitude." *KdrV* A183.
[564] As Patricia Kitcher notes, Kant's doctrine of apperception is a promotion of the unity of the self against the Humean bundle theory of the self, which reduces the self into a flux of perceptions that does not pertain to a single self-consciousness across time. Kitcher (1993): 99.
[565] *Diss* 2:407.

actually integrate the event-causal series into an underlying mode of activity that can serve as the metaphysical basis of character, in virtue of which passing events of the agent's psychology are subject to oversight through the agent's character.

Of course, the idea of a supersensible mode of interaction that underlies the spatiotemporal framework of sensibility, for the Critical Kant, is not an object of theoretical insight. What happens, though, is that transcendental idealism can properly orient the cognizable events in the spatiotemporal framework as metaphysically derivative, in response to the antinomies generated by reason's clash with a dogmatic realism about this framework. This opens up a space for practical reason to cognize itself as supersensibly active in the form of absolute freedom, as Kant claims it can, in *KdpV*, by implication of the fact of the moral law.[566]

Neo-Humean Contemporary Libertarianism and the Regressive Method

The obvious objection to these concerns about the Humean world, in which we have only particular events that do not causally constrain one another, is that the Humean world has the virtue of having complete proportion between its metaphysical claims and empirical evidence.[567] We can build the Humean world outwards from epistemological primitives, impressions, which provide immediate non-inferential knowledge. On the

[566] *KdpV* 5:4. There are realist and anti-realist interpretations of practical cognition, i.e. *Glaube*, which Ostaric discusses in Ostaric (2017).
[567] Eric Watkins acknowledges this virtue of the reductionist metaphysics of David Lewis, in which all macro-level phenomena are reduced to our most basic impressions, in Eric Watkins (2005): 398.

other hand, Kantian metaphysics, involving both schematized and unschematized contingency relationships, clearly introduce a metaphysics that transcends any aggregation of simple empirical impressions.

The Humean empiricist is scrupulously committed to impressions in its construction of ontology, and this methodology leads to devastating consequences for the metaphysics of the self. Substance for Hume is a psychological, not an ontological reality, that hypostasizes what is in reality a succession of impressions, that in its consistency gives an *appearance* of substantiality. A rigorous empiricism can provide no grounding for the ontology of substance-hood, and so an idea of an underlying unity of the self binding together these different impressions has to be dismissed.[568] The Humean self is rather a disconnected stream of interior impressions, none of which provide a basis for an identical and stable self.[569]

The empirically based theories of free will in contemporary libertarianism, which focus on events grounded in simple impressions and deny any underlying metaphysical substructures, are Neo-Humean. The key methodological principle of Humean empiricism is the copy principle, in virtue of which terms are defined by way of empirical impressions to which they correspond. Those ideas that lack empirical correspondence are dismissed as meaningless.[570] This methodology leads to stunted theories of freedom that present it as an indeterministic event or uncaused powers. The absence of character leaves the agent a spectator on her own psychology. An ontology grounded on simple

[568] See Hume's *Treatise* (1949): 220.
[569] This changing stream for Kant is *empirical* apperception, which for him presupposes a *transcendental* apperception. See *KdrV* A107.
[570] See the formulation of the copy principle in Hume (1999) 7.1.2

impressions cannot provide the unity of the self in which event-causal series do not simply fall out by chance or line up with pre-existing series outside of the agent, but which are subject to the agent's second-order endorsement through character.[571]

ECL theorists like Balaguer adopt a metaphysics of mind-brain materialism in which there is a token-token identity between neural events in the brain and action.[572] Any metaphysics of character is dispensed with, as ontological excess that is not deducible from simple impressions. A materialist metaphysics, though, in which neural events are ontologically fundamental, forces us to articulate free will in terms of events that either arise by chance in probabilistic causation or through deterministic causation. A materialist metaphysics eviscerates the possibility of a genuinely active power of spontaneous causality, since it forces us to articulate it in terms of mechanical causation at the microphysical level that only passively follows from its context.[573] Reason in this empiricist metaphysics is not self-determining, but only on the same level as the event-causal order, and can only work on behalf of psychological forces that contextualize it. Hence, Humean reason is a "slave to the passions."[574] Neo-Humean empiricism assumes that the causality of freedom is a property of impressions, and does not lie in schematized pure concepts.[575]

---

[571] I am thinking of Harry Frankfurt, who saw the human person as not only willing, but willing to will. See Frankfurt (1971): 5-20.
[572] Mark Balaguer (2010): 70-1. Kant rejects mind-brain materialism by way of the unity of consciousness: each thought is united under the structure of the "I think," but a purely materialist mind would distribute different thoughts to different brain states. See the discussion in Henry Allison (1989): 195.
[573] This materialistic metaphysics amounts to an "objective view of the world," which considers the mind apart from the first-person perspective of the "I think." See Allison (1995): 350.
[574] Hume (1949): 2.3.1.2.
[575] *KdrV* A137-8.

Balaguer explicitly adopts an anti-metaphysical stance in which the question of free will is an empirical one resolvable by observation. For Balaguer, metaphysical questions that are not reducible to an empirical question resolvable through science, or through a strictly logical procedure, are in fact meaningless.[576] Kane, another ECL theorist, also pledges, in an effort to put theories of free will in conversation with modern science, to not make any ontological commitment in a metaphysics of agency that is not something to which a compatibilist would appeal. This principle restricts Kane from positing any special and irreducible genus of causation that goes beyond the event-causal order a compatibilist would need to articulate a theory of freedom.[577]

There are some who would find a form of freedom based solely on impressions, and so detached from the unity of the self, the core of a promising kind of compatibilism. An agent subject to event-causation in the form of a deterministic series stretching beyond itself, and so possessing a practical reason subject to outside forces, could nevertheless have liberty of spontaneity[578], i.e. an ability to act according to her will without constraint. An agent subject to external causation could nevertheless choose to do what she is doing, in the sense that she is the immediate (though not ultimate) cause of her actions. Such an agent could also have hypothetical liberty, in virtue of which, if she chooses otherwise, she would act differently.[579] The introduction of indeterminism, as we see in contemporary libertarianism, could further secure a liberty of indifference, by which the agent is not necessitated to act by prior causes.[580] This agent also could act

[576] Balaguer (2010):22-24.
[577] Kane (1996): 115-123. Manuel Vargas also recommends a commitment to naturalistic plausibility, cited in Timpe (2013): 46.
[578] Hume uses this phrase in *Treatise* (1949): 2.3.2.1.
[579] Hypothetical liberty comes up in Hume (1999): 8.23.
[580] Hume uses the phrase "liberty of indifference" in Hume (1949): 2.3.2.1.

according to what it consciously chooses, and also have the hypothetical liberty to act differently if its reasons happen to be different. Neither the compatibilist nor the libertarian agent would require any metaphysics of substance in virtue of which agent is able to exercise some power that transcends the event-causal order.

However, both of these agents suffer from the same defect of a lack of independence from the event-causal order. The compatibilist agent is controlled by prior events, while the indeterministic libertarian agent is controlled by whatever the chance process happens to yield. The turnspit freedom Kant discusses in *KdpV*[581] is unsatisfactory because its will is directed entirely by processes outside of its control. To provide the kind of self-determining ultimacy that goes beyond turnspit freedom, we need the metaphysics of substance-causation that is not answerable to the event-causal context, whether it is indeterministic or deterministic.

To go beyond turnspit freedom, in which practical reason is on the same plane as contextual psychological causes, we need to break free from the epistemological methodology that restricts us from positing substance-causes. We can turn to Kant to break free from the neo-Humean empiricism of contemporary libertarianism, in establishing a metaphysics of substance in virtue of a priori structures of the understanding that make empirical cognition possible. We can appeal to Kant to show how empirical knowledge of neural events and the social sciences, the epistemic prestige of which ECL theorists presuppose, cannot be accounted for solely in terms of Humean impressions.

[581] *KdpV* 5:97.

The Kant of Karl Ameriks adopts a regressive method, in virtue of which the acceptance of the reality of objectively valid empirical knowledge requires a priori structures of the understanding. The regressive method then moves towards the appearance/thing in itself distinction of transcendental idealism, by establishing that reason cannot posit its unconditioned aspirations within the categorical structures of the understanding.[582] Kant's regressive method, by which the empiricist's rejection of metaphysical structures of experience is countered by a demonstration of the necessity of a priori structures to empirical knowledge, occurs most clearly in Kant's Transcendental Deduction. This is a very contentious section of Kant's work, but it will suffice here to appeal to it to at least point the way out of the inadequacies of Humean empiricism.[583] Empirical knowledge cannot consist in isolated and simple impressions, but instead involves a synthesis of different impressions in a judgment. The mere sensation of redness, for example, cannot itself serve as the basis of knowledge: instead, redness must be connected to an object (e.g. the apple is red) or at the very least assimilated to a concept by which we can repeatedly recognize redness in different experiences. Knowledge cannot consist in the mere passive reception of impressions, but must involve a spontaneous synthesizing capacity of the mind. Individual impressions are assimilated to the form of temporality and given a place in it. For a manifold of different impressions to constitute knowledge of an object, this manifold must be unified across time. Otherwise, the manifold will degenerate into disconnected parts—a rhapsody of

[582] Karl Ameriks (2003): 5. Ameriks opposes a *progressive* interpretation of Kant's method, which is not an attempt to establish a prior conditions of empirical knowledge, but which attempts to move from subjective representations to knowledge of the external world—it is anti-skeptical. Ameriks (2003): 51. For instance, Strawson thinks that Kant in the Transcendental Deduction attempts to prove, rather than presuppose, knowledge of the objective world through the reality of self-consciousness.
[583] Kant wrote two versions of the Deduction, an A version (*KdrV* A95) and a B version (*KdrV* B130).

perceptions.[584] Kant defines an object as "that in the concept of which the manifold of a given intuition is united."[585] Empirical knowledge requires some consistency across time, since a lawless empirical order would threaten the possibility of non-reproducibility of impressions, e.g. cinnabar could be red one day, black another, etc. Mere impressions cannot ground the necessity essential to knowledge.[586] A substance-metaphysics is needed to unify one's experience across time, since the mere formal framework of time cannot itself be a sensation that accounts for the sequence of sensations.[587] Empirical knowledge of objects requires a stable whole—i.e. a substance-metaphysics—of which changing events/properties are characteristics. A certain unity must characterize even basic counting, insofar as a complete isolation of individual impressions would not allow the accumulation necessary for counting.[588] Empirical objects need to be understood as extensive magnitudes, which are not discrete impressions but united aggregates.[589] The cognition of water freezing cannot be based solely on discrete impressions, but instead requires an assimilation to the structure of temporality, and an objective asymmetry in the event-relationships (not in their mere perceptions) based on a causal law by which successive events are united in a cause-and-effect relationship.[590] If this causal relationship is to be universally valid knowledge, it must rest on a priori structures and

[584] *KdrV* A156/B195.
[585] *KdrV* B137.
[586] Kant also shows how a priori categories are needed to move from judgments of perception to judgments of experience in *P* 4:298.
[587] "Since that in which sensations alone can be ordered…cannot itself be another sensation…" *KdrV* A20/B34. Sensations do not come with a label indicating their temporal position. Guyer (1987): 170.
[588] *KdrV* A103.
[589] *KdrV* A162/B203.
[590] *KdrV* B162. Guyer cites Graham Bird to note that differentiating one event from another requires a causal law by which states have a necessary order. Guyer (1987): 249.

not mere sensation.[591] Alteration requires a permanent substantial substructure. Otherwise, it is not alteration but merely disconnected impressions that may be only subjectively successive but objectively coexistent.[592] Alteration requires *some thing* undergoing alteration.

A key footnote in Kant reveals that his delineation of the a priori structures of experience, particularly his principles of time-determination in the Analogies of Experience, are not merely epistemological principles about how we come to know something, but indicate also ontological conditions of cognition. Kant notes that the substantial structures and causal laws of the Analogies, whereby time-determination of events is shown to be constituted in terms of mutually interactive substances subject to causal laws, allow us to make an inference from merely ideal relations in thought to real relations in things.[593] That is, the Analogies of Experience allow us to transition from a merely ideal connection in thought to a real connection in objects: we are able to non-arbitrarily identify two perceptions as simultaneous through an objective grounding in substances in mutual interaction (third analogy), as successive through objective grounding in a causal law (second analogy), as successive or simultaneous against an objective grounding in persisting substances that represents the permanence of time. The move in the Analogies is therefore from an ideal relation in thought of two events to its real relation in objects through a certain metaphysical constitution. The ideal/real distinction comes from *ID*, where Kant rejects the idea that the unity of the world can be

[591] See Guyer (1987): 126 for confirmatory quotes. See also *Prol.* 28 for Kant's claim that no causal knowledge can occur but for a relationship to the categories.
[592] *KdrV* A187/B230.
[593] The relevant footnote is *KdrV* A218/B265.

a mere ideal unity based on thought, as in Leibnizian pre-established harmony, and not on real causal interaction.[594] The ontological basis underlying the cognition of time-determination[595] points the way out of neo-Humean accounts of the will based on de-contextualized events, and towards a unified and controlled agent capable of being a participatory ground of its own psychological events.

Unity of the self, rather than being something an empiricist can reject, actually is presupposed by empirical knowledge. Kant's doctrine of apperception is the idea that a unity of self-consciousness—an "I think"—must underly empirical knowledge, because empirical knowledge requires a subject capable of representing it as knowledge.[596] All our empirical knowledge must be in space and time. Time and space themselves are unities, insofar as we have cognition in one universe in the same spatiotemporal structure, and do not transition to parallel universes with each new experience. The unity of time and space cannot come from the senses alone, but can only come from the categories, since the discrete impressions of the senses cannot unite themselves in a spatiotemporal structure. Kant in the First Analogy argues for a substance metaphysics necessarily underlying changing events, since an enduring substance grounds the unity of time by forming the necessary contrast between what stays the same—i.e. time—and changing determinations of this abiding substructure.[597] All cognition of empirical objects must rest on the a priori structure of the understanding, and fall under the apperception applying

[594] *ID* 2:390.
[595] Watkins identifies this ontological element of the Analogies through a survey of pre-Critical debates over the metaphysics of causality, and so pushes back against merely epistemological accounts of the Analogies that identify them as offering only justificatory relationships between judgments. Watkins (2005): 200.
[596] *KdrV* B132.
[597] *KdrV* A188/B231.

those categories.[598] The necessity of substance-metaphysics for knowledge of objects, and not mere strings of subjective modifications, can lead us out of a substance-less agency based only on events.

The transcendental doctrine of apperception, in which we connect different impressions into a single self-consciousness, actually mirrors Kant's model of causality by which an agent-substance can exercise control over its own action. The connective faculty of apperception is a continuous activity that is not identifiable with a temporally determinate event. Apperception is a non-empirical ground of our awareness of objects, insofar as we do not have a direct intuition of the self.[599] The continuous and temporally indeterminate activity of apperception, which serves as the ground of the connectivity of representations even while not being a spatiotemporal event itself, mirrors the kind of substantial activity by which an agent exercises control over its empirical character without manifesting itself as an antecedent ground in time.[600]

An important distinction at this point must be made between those who challenge the empiricist theory of free will solely in terms of an epistemic spontaneity, by which the receptivity of our sensibility is complemented by a spontaneous structure of taking intuitions under certain concepts as well as the "I think," and those who challenge the empiricist theories in terms of a genuine unconditioned causal power. The spontaneity of our understanding, by which it prescribes the connective tissue of empirical knowledge, provides the ground in Allison for the deliberative standpoint of the intelligible character. Allison provides a complement to the event-causal order emphasized by the empiricist by

[598] *KdrV* B161.
[599] *KdrV* B407.
[600] Watkins ties transcendental apperception to Kant's model of causality in Watkins (2005): 272.

way of a spontaneity of judgment in the practical domain that takes certain inclinations as appropriate bases of action.[601] Allison accounts for Kant's dissatisfaction with compatibilism not by pointing to some belief in an unconditioned and non-empirical power of agency, but by noting that, for Kant, epistemology involves a self-consciousness that cannot be accounted for in terms of receptivity to external impressions. "Yet the human being, who is otherwise acquainted with the whole of nature solely through sense, knows himself also through pure apperception, and indeed in actions and inner determinations which cannot be accounted at all among impressions of sense…"[602]

But, this merely epistemic form of spontaneity is not itself a form of non-empirical causality, and so the Allisonian view of freedom does not deliver the agent from subjection to the event-causal order. Epistemic spontaneity is compatible with a metaphysics in which the agent is a mere accident of larger and impersonal causal networks, e.g. the Spinozistic agent that is mere accident of the will of God. Kant was against the kind of dogmatic monism that assimilates the agent to its larger causal background. The Spinozistic agent may have its own thoughts—i.e. it may be conscious of itself as an spontaneous subject—but because of the exclusive substantiality of God, these thoughts are ultimately manifestations of God through the human mind.[603]

The regressive method, as I hinted at the opening of this section, is a multi-stepped process in which the establishment of a priori categories, when tested against reason's aspiration for the unconditioned, leads to the appearance/thing in itself

[601] Allison (1990): 39.
[602] *KdrV* A547/B575.
[603] See Karl Ameriks article "Idealism and Kantian Persons," in Ameriks (2012): 129.

distinction of TI. While the formal element of self-consciousness is not transcendental freedom—the Paralogisms deny any inference from the formal elements of self-consciousness to a substantial self[604]—the spontaneity of the combinatory role of the understanding is part of a larger Kantian system. From the spontaneity of the understanding, and the givenness mediated by the sensibility, we arrive at an appearance/thing in itself distinction. The formal categorical structure of self-consciousness is in itself empty apart from some content given from a source independent of it.[605] The way something appears to self-consciousness, then, can be distinguished from the intrinsic nature of what is given (i.e. by thinking of it independently of its givenness). The spontaneity of the understanding, though, is not content to merely combine immediate sensations in a localized context, since human reason by nature tries to extend itself in a totalizing way towards the conception of the world in its unconditioned limits. Reason's striving for the world as a totality allows us to distinguish, in turn, between an intelligible world—of which we can only think—and the sensible world of immediate experience. The necessity of intelligible world is solidified in virtue

[604] The Paralogisms begin at *KdrV* A341/B399. It needs to be noted that Beatrice Longuenesse has argued that the metaphysics of personhood does not require a substantial soul, but a merely formal structure of a numerically identical consciousness uniting different moments in time. Longuenesse thinks that Kant's later deduction of absolute freedom from the existence of the moral law represents an inconsistency in his rejection of the arguments of the rational psychologist. Beatrice Longuenesse (2017): 156.

[605] I am rejecting here bowdlerized or naturalistic accounts of Kant's transcendental psychology, such as that of Patricia Kitcher, who presents Kantian consciousness as a causally connected system, without any appeal to the spontaneity of the "I think." Kitcher's Kant presents mental states as causally connected in us, but not for us, and so is a proto-functionalist account. See Allison's critique in Henry Allison (1995): 341.

of the antinomies, which show the contradictions that arise when reason tries to think the unconditioned in terms of the sensible world.[606]

Reason's demand for the unconditioned thus relativizes the domain of schematized categories in empirical cognition. We therefore have the basis of the hybrid empirical-intelligible agent, which has an intelligible character as a real metaphysical causality and not a mere epistemic perspective.

The Allisonian picture contrasts between what is given to self-consciousness and the spontaneity of self-consciousness itself without at the same time permitting the distinction between the way self-consciousness takes up its experience and the thing given to be an ontological distinction between kinds of thing. If spontaneity is a mere mode of conceiving of experience, and the controlling ontology is givenness to the mind, then spontaneity appears to be a merely formal category, arrived at through an abstract sort of thought experiment, that lacks any genuine relevance to cognition. Moreover, the Allisonian emphasis on epistemic spontaneity of the understanding neglects the tendency of reason to seek beyond the immediate grasp of the present moment towards the ultimate conditions of things. Reason striving towards the unconditioned leads to the antinomial conflict with the conditions of the understanding that forces us to adopt a belief in the non-empirical power that Allison's non-metaphysical view of freedom rejects. Allisonian epistemic spontaneity fails to fully embrace the regressive method, which not only posits a priori structures of cognition but also makes a real metaphysical distinction

[606] The progression from the spontaneity of the understanding and the duality of the understanding and sensibility, to the ultimate ontological framework of two worlds, is echoed in Dieter Henrich's piece, "Freedom as the 'Keystone' to the Vault of Reason," in Dieter Henrich (2003): 46.

between what appears through these structures and what is given to them in virtue of antinomies generated by reason's aspiration towards the unconditioned.

Conclusion

In this chapter, I explored three kinds of contemporary libertarianism from a number of different angles. I considered the importance of character in an account of agency for securing imputability and rationality. I considered how the metaphysics of transcendental idealism are best suited to do this, since this metaphysics can lift one out of a mechanistic sequence of events in which the agent is a disappeared spectator on its own behavior. I considered the issue of the unity of the world with respect to causally disconnected empirical fragments. Finally, I considered the merits of a regressive strategy, that seeks to account for empirical knowledge in terms of rationalist a priori structures, for delivering us from limited empiricist conceptions of free will.

As for Longuenesse's contention that an identity of self-consciousness underlying different impressions (see footnote 92), evident in the original unity of apperception established in the Transcendental Deduction, is sufficient for personhood, this claim is subject to the same objections made against NCL. That is, apperception makes the self the mere subject of mental events, without giving the self any causal role in shaping those events. Apperception is merely epistemic connective tissue, as is the spontaneity of the categories, in virtue of which discrete subjective representations becomes knowledge of objects.[607] Apperception can give the agent a deliberative *perspective*, in which exercises spontaneity by taking inclinations as consistent with a certain maxim, and is not merely

[607] See the discussion in the *Prolegomena* of the transition from judgments of perception to judgments of experience, in *P* Sections 18-19.

controlled—at least not phenomenologically—by its strongest desire. However, agency that is a mere mode of self-conception cannot exercise any real causality over its strongest desires. Persons that merely observe their own mental states as their formal containers are not appropriate subjects of Strawsonian reactive attitudes.[608] Their actions happen to them, because their freedom is not a form of causality, even though they can "conceive of themselves" as exercising deliberative agency in light of norms.

[608] See the discussion of exemptions to reactive attitudes of praise and blame in Gary Watson (1993): 134. I evaluate the standpoint non-metaphysical theory of freedom at greater length in Chapter 3.

CHAPTER 5

AGENT-CAUSATION WITHIN THE FRAMEWORK OF TRANSCENDENTAL IDEALISM

The problems of contemporary libertarianism will lead to compatibilism or a free will skepticism if a new metaphysical framework is not provided. Agent-causation may provide the sort of control that event-causal and non-causal theories cannot provide, as agent-causation directly involves the agent in action instead of delegating it to mere sequences of events. However, it is not enough to merely assert that agent-causation exists, without explaining the mechanism by which an agent is to directly bring about action. Claiming that an agent directly causes an action, as a substance, and not through events, invites charges of arbitrariness, since removing mental states of causal involvement seems to preclude rationality.[609] It will not do, moreover, to merely assert that the agent-cause does not directly cause the action, but instead causes itself to act on certain reasons and not others, as Randolph Clarke does.[610] A metaphysical framework has to be developed to show exactly how such structural causation[611] is supposed to work.

In section one of this chapter, I will identify features of the metaphysical framework I will use to develop an account of agent-causal freedom. In section two, I will discuss advantages of the metaphysical framework of agent-causation I am using. In

[609] Davidson, for instance, articulates rationality in terms of a causal explanation linking certain states of the agent to action. Donald Davidson (2001): 3.
[610] Randolph Clarke (1993): 194.
[611] A structural cause is a cause that accounts for why two events are linked. For example, the wiring in the doorbell accounts for why pressing the button of the doorbell causes the bell to ring. See Fred Dretske (1988): Chapter 2.

section three, I will address objections, in historical Kant scholarship as well as contemporary commentary on agent-causal theories, towards my agent-causal account. First, I will address the objection that the immaterialist metaphysics I am proposing permits itself to slide into a dogmatic metaphysics the Critical philosophy is supposed to repudiate by restricting application of the categories to sensible intuition.[612] The second objection I will address has to do with the supposed incoherence of the notion of a timeless agent.[613]

Key Features of the New Metaphysical Framework of Agent-Causation

In chapter 4, I identified the event-event ontological structure of contemporary libertarianism. Even O'Connor accepts that agent-causation, which has a power of self-determination that distinguishes it from the mere passive transferences involved in event-causation, is causally united with the event-causal order.[614] What happens with a metaphysics of agency based entirely on events is that the agent disappears. An agent subsumed in deterministic events acts according to what is already there[615], i.e. what is already built into prior causes beyond the agent's control. The indeterministic agent is subject to whatever happens to be the result of a chance process. These agents may be able to act, in the sense that their acting on their desires is externally unconstrained, but

[612] For problems related to articulating agent-causation in terms of contemporary science, see Derk Pereboom and Michael Mckenna (2016): 251.
[613] Henry Allison makes this objection very strongly in Henry Allison (1990): 51.
[614] O'Connor makes this distinction between agent and event-causation in Timothy O'Connor (2000): 72. O'Connor is a nonreductive materialist. Pereboom (2001): 69.
[615] McKenna uses this language to describe the determined agent in Michael McKenna (2008): 195.

they are not able to *intentionally* act, since factors apart from their intention control them.[616] There is no possible world for these agents, given an identical past, in which they can act otherwise than they did, because they lack any contra-causal power apart from events in their past.[617] Metaphysical grounds, that are not themselves events, contextualize events in a way that subjects them to agential control. The metaphysics I will invoke differs from the event-causal order of contemporary libertarianism in the following ways: it is 1) tiered, 2) a whole and not an aggregate, 3) has diversified contingency relationships, and 4) does not involve either causal inertness or overdetermination

*Tiered metaphysics***.** The metaphysical framework of transcendental idealism I use in framing agent-causation is tiered, insofar as reality is understood not only laterally, in terms of events located along the time-series, but in terms of fundamental-derivative relationships. The lateral time-series of events is seen in TI in terms of schematized versions of the pure categories. Schematization applies the pure structure of thought through a transcendental time-determination to a particular intuition subject to the temporality of inner sense.[618] The process of schematization sets up tiered metaphysics. There is the level of schematized categories in space and time, which have to do with a contingent form of sensibility that does not represent the nature of absolute reality. There is also the level of the intellectual synthesis of the B-Deduction, in which we think of

[616] Mele makes the distinction between an ability to act, and an ability to act intentionally, in Mele (2006): 19.
[617] Without such a genuine contra-causal power, an agent resembles someone subject to direct neural intervention, as in Derk Pereboom's manipulation thought experiments, cited in Mele (2006): 139.
[618] Kant discusses the process of schematization at *KdrV* A137/B176. Guyer notes that the schematism shows how the categories can be instantiated in space and time. Guyer (1987): 158.

reality in general in terms of the pure categories.[619] Things in general represent an absolute level, that exists determinately apart from any subject, whereas spatiotemporal appearances represent a level that is in itself indeterminate and only becomes determinate with respect to human sensibility. The contrast in determinacy between things in themselves and appearances can also be seen in the understanding of motion in *Metaphysical Foundations*: Kant claims here that motion of matter in space does not have its power to communicate motion intrinsically, but only in virtue of more fundamental forces that cause it to occupy space even prior to any motion of its own.[620] The superficially apparent motion of matter is, therefore, indeterminate.

*A whole and not an aggregate***.** Unstructured events, without an underlying intelligible framework, run the risk of an empirical chaos, in which a meaningless string of events form a random aggregation instead of a systematic whole. Rather than a mere unstructured sequence of events, the Kantian metaphysics I am developing situates empirical events in an intelligible order of an immaterialist metaphysics. The intelligible order serves as the appropriate site of the unconditioned, in virtue of which reason can strive towards the idea of a systematic unity by which empirical chaos is replaced by empirical lawfulness.[621] Kantian metaphysics complements the horizontal relationship between empirically conditioned events, which is otherwise an unsystematized

[619] Kant discusses the *synthesis intellectualis* at *KdrV* B151.
[620] *Foundations* 4:537. Kant privileges in this work the metaphysical-dynamical mode of explanation over the mathematical-mechanical mode of explanation. The former invokes fundamental forces at the basis of matter, whereas the latter invokes nothing more than the shapes and motions of particles moving through emptiness. Ibid., 4: 525. Another way of phrasing this is that extensive magnitude presupposes intensive magnitude. Guyer (1987): 201.
[621] Ostaric notes how systematic unity is not a mere heuristic fiction, i.e. an optional stance needed only to maximize our empirical knowledge, but is in fact a *transcendental* principle posited as objectively valid. Ostaric (2009): 164. See *KdrV* A653/B681.

aggregation, with an unconditioned principle in virtue of which empirically conditioned events have an inner coherence.[622] The specter of empirical chaos is banished in virtue of the principle of purposiveness, pertinent to the faculty of judgment, in virtue of which we can subsume particular intuitions under general principles.[623] Judgment's principle of purposiveness mediates between reason and the understanding, and so it brings the empirical cognitions of the understanding under the aim of completeness in the unconditioned of reason.[624] A materialist, event-based metaphysics lacks grounds which would otherwise be able to account for the sequence of events, and unify them according to the rational principle of the unconditioned. In addition to the unconditioned, pure apperception is also integral to bringing unity to the aggregate of empirical events: it is through complementing empirical events with pure apperception that they are united in a single consciousness. In a purely materialist metaphysics without such structures, thoughts are identified with neural events in different parts of the brain, and so thought lacks unity.

A whole structured according to the rationalist principles of the unconditioned, as well as that of transcendental apperception, has a significant advantage over the mere aggregation we find in contemporary libertarianism, in which actions unfold according to chance processes that lack a sufficient cause in the agent's motivations. The integration of the empirical and intelligible levels lifts us out of the randomized event-causal sequences towards an intelligible character that forms maxims that structure its

[622] Ostaric makes this point in Ostaric (2009): 156. Three principles inform the systematicity of nature: homogeneity (similarities allow us to identify common categories), specification (fine-tuned awareness of differences), and continuity (no empty transitions from specific to general). *KdrV* A657/B685.
[623] For an excellent account of this idea, see Henry Allison (2001): 38.
[624] *KdU* 5:168. Guyer wishes to suppress judgment's mediation between freedom and nature, as presupposing a dogmatic metaphysics, in Guyer (1979): 54.

participation in the event-causal order. In event-causal libertarianism, we articulate action merely in terms of a sequence of psychological events leading up to action. This sequence represents a mechanical empirical series, instead of an integrated world systematically structured according to the rational principle of the unconditioned. Delegating action to a series of events sacrifices rationality, as the agent is subject to the events and is not able to exercise control over them. For example, if someone jumps because an unseen but nearby person says "Boo!", there may be a sequence of desires and beliefs leading up to the action, but one can hardly describe the action as rational.[625] It is more of a blind reflex.[626] Antecedent events need to be taken up by one's character in order to truly have a rational significance. The only way to rescue an agent from this merely reactive stance of subjection to psychological events is by permitting it an unconditioned power of causality in a world in which intelligible and empirical structures are systematically united.[627]

A desire for chocolate arises in one person, and that person goes to the store to buy some. Another person gets the same desire, but refrains from going. Assigning an intelligible ground for these sequences of events allow us to account for diversity in outcomes in terms of maxims, and so we can readily amend the event-causal accounts that attribute diversity in outcome to indeterministic process that are incompatible with

[625] Here, I am rejecting certain compatibilist accounts of rationality, in which an agents acts on reasons but is ultimately determined by non-rational pressures. For an example of this, see the fourth case in Pereboom's manipulation arguments in Pereboom (2001): 115. This rejection echoes Kant's rejection of compatibilist rationality, which I discussed in Chapter 2.

[626] This example, and a larger discussion of the difference between reasons and causes, can be found in Elizabeth Anscombe (1957): 23.

[627] The alternative, of an agent whose actions are events that happen to the agent, is the sort of pathological actor that Strawson says would not be the proper object of reactive attitudes—e.g. someone who is profoundly intoxicated. See the discussion in Watson (1993): 123.

agential control. Instead of conceiving of action as a linear sequences of events that are ungrounded in the agent's character, we can account for the outcomes in terms of maxims, and so event fragments at the empirical level become intelligibly grounded in a unified agent.

Diversified contingency relationships. The contrast between the schematized category in space-time and the pure category applied to things in general in the intellectual synthesis allows for a diversification of the contingency relationship. Rather than merely events in time that are contingent insofar as they depend on a previous event, the intellectual synthesis allows for a form of contingency that is not in time but involves only the recognition that something could have been otherwise.[628] Diversified contingency relationships allow us to form a metaphysics for genuine agency that goes beyond false agent of ECL: if an agent acts in ECL because a behavior was caused by a psychological event, one can legitimately ask in virtue of what does the psychological event also represent an act, and not a mere internal push to which the agent is subject? Introducing another mediating event will only expand into an infinite regress, since we have to ask in virtue of what does the mediating event represent an act?[629] An intelligible form of contingency can represent a ground of the empirically contingent series, so that the agent still participates in the event-causal order, but nevertheless shapes this order according to maxims that are *intelligibly* contingent. An intelligible ground does not

[628] Kant distinguishes between empirical and intelligible contingency in the remark on the thesis argument of the fourth Antinomy. *KdrV* A460/B488.

[629] Gary Watson shows how the libertarian presupposition that, to be responsible for an act, we have to be responsible for the kind of self that produces it, leads to an infinite regress. This responsibility for the self can only occur, he says, if the self both transcends and intervenes in the causal nexus, as both the author and the producer of its action. Intelligible and empirical forms of contingency can provide the basis for this transcendent and intervening self, insofar as the empirical character acts in time while the intelligible character serves as the self-determining ground outside of time. Watson (1993): 143.

require an antecedent event in order to be a genuine act, since it is not in the time-series of *empirical* contingency, and so the prospect of an infinite regress does not threaten the account.

For Kant, the self bridges two worlds.[630] On the one hand, it is part of the event-causal order in the empirical world, in which its psychological events are subjected to antecedent conditions. At the same time, the immaterialist Kantian metaphysics allows the self to exercise an unschematized form of causality that does not manifest itself as an event in time. This form of intelligible contingency functions in the Incorporation Thesis[631]: the empirical motivations manifesting as events in time are not themselves sufficient to determine an agent's behavior, but must first be incorporated into a normative principle.[632] The grounding of empirical sequences in maxims means that the intelligible character, in virtue of maxims, privileges certain reasons for action over others. There is a combination, then, of a deliberate selection of maxims with automatic processes of event-causation at the empirical level. In a similar way, cognitive scientists identify automatic activation systems involving muscle contraction and limb trajectory over which we lack conscious control, while at the same time the agent consciously chooses a certain action.[633]

The intelligible activity of the self does not figure as an event in time—e.g. first there is a maxim, then there are desires, then a belief, then action—but instead acts as a structural feature that controls the flow of events. In this way, my account of agent-

[630] Dieter Henrich (2003): 52.
[631] The Incorporation Thesis comes from Henry Allison in Henry Allison (1990): 5.
[632] Kant defines a maxim as a subjective principle of practical reason. Our maxims mediate between the moral law and our empirical motivations. *Gr* 4:402.
[633] See the discussion in Timothy O'Connor (2000): 123.

causation is similar to Randolph Clarke's, who posits the agent as a cause that endures throughout the action-sequence, not as an immediate cause, but as the basis for why certain events, taking the form of reasons, cause certain actions.[634] Diversified contingency relationships are essential to explaining how the agent functions as an enduring basis for linkages at the empirical level. If all we have is contingency in the form of antecedent events, we cannot articulate the agent as an underlying condition linking psychological events and actions, since the underlying condition stretches across different moments in time. A disadvantage for Clarke is that the agent-cause is not supposed to contribute to scientific explanation of the agent's behavior, since the event-causal order for him is the fully self-standing basis for such explanations.[635] Agent-causation has moral significance for Clarke, not scientific significance. Diversified contingency relationships are necessary to rescue the moral significance of agent-causation from an illusory status. If event-causation is all there is, the moral view is defeasible, and we may as well treat one another as purely reactive machines (e.g. turnspits). With diversified contingency, we can view the empirical series as a contingent appearance dependent on the metaphysically fundamental intelligible order, in a way that genuinely integrates the agent-cause into the metaphysics of agency rather than making it into a superfluous feature taken up only in moral discourse.

Diversified contingency relationships also permit us to rebut certain objections that point to the lawfulness of the empirical order to refute agent-causation. If an agent-causation existed, the objection holds, we should see frequent disruptions of the

[634] Randolph Clarke (1993): 194.
[635] Ibid., 199.

lawfulness of the empirical order, as the agent freely selects its behavior in ways that are unconstrained by laws of nature.[636] Barring some sort of wildly consistent coincidences between what the agent-cause selects and the empirical order, the fact that there is no empirical deviation from lawfulness suggests there are no agent-causes.

But, again, this objection presupposes a monolithic notion of contingency that forces the agent-cause into the empirical time-series as a disruptive force. Once one has incorporated transcendental idealism, one can form a view in which the lawfulness of empirical series is unbroken, and yet the agent-cause can still exist as an underlying ground of this series. It is not as though the agent-cause has to emerge in time and steer the series of events in a self-determining way, leading to violations of empirical laws.[637] Instead, diversified contingency relationships can allow us to understand how an agent can serve as the very basis of empirical laws, since it provides a continuous substructure that accounts for the linkages of certain psychological events and actions. As the basis of empirical laws, the agent is not subject to them, and so, at the intelligible level, the principle that alternative possibilities are required for free action can be met.

An agent-cause does not have to participate in the empirical-series, so long as we do not view the empirical-series as metaphysically absolute. For the Kantian, matter is merely phenomenal substance, and not an object of pure understanding characterizing reality in itself.[638] Instead, the agent-cause can serve as a constitutive ground of the empirical series. There is an empirical contingency involving precedence in time—Brutus

[636] Derk Pereboom develops such an evidence-based objection to agent-causation in Derk Pereboom (2001): 79. Pereboom applies this objection to both deterministic and indeterministic series.
[637] Some versions of compatibilism actually appeal to a miraculous ability to break laws of nature. Timpe (2013): 72. My agents do not break laws, but ground them.
[638] *KdrV* A277/B333.

stabbed Caesar and caused Caesar's subsequent death—but there is also a constitutive contingency without which an object cannot appear as it does—without electrons, there would be no lightning.[639] The constitutive ground does not precede what derives from it in time, but constitutes its inner nature. The lawful empirical series can be understood as the manifestation of the freedom of the agent-cause, which acts as an inner ground of the empirical series.

We can appeal to the notion of a substantial form to explain the intelligible-empirical relationship: the intelligible structure imparts a form to the empirical domain insofar as it determines how its contents are arranged into laws with certain relata. The relationship between the intelligible and empirical structures, then, can be likened to the hylomorphic relationship we find between soul and body in Aristotle, in virtue of which the soul provides a structure to the matter of the body, which in itself is merely shapeless potential.[640] The structuring teleology at the intelligible level comes in the form of maxims that originate not in time, but in rational representation.[641] The character at the intelligible level is the site of a meta-reason, i.e. a fundamental orientation of the will in virtue of which individual acts in time take on a certain tendency, and so the intelligible character creates the causal patterns not as an event itself but as a bracketing structure.[642] The substance of the self uniting its spatiotemporal aspect is at the intelligible level,

[639] The lightning/electron example does not perfectly capture the relationship between the empirical and the intelligible in Kant. Electrons, though more fundamental and absolute than the lightning, themselves occupy space and time. Electrons count, therefore, as the *empirical* thing in itself. *KdrV* A45/B62.
[640] Aristotle, *De Anima*, ii.1.412b6-9. In Leibniz, too, we are exposed to a rehabilitation of the Aristotelian idea of a substantial form. Leibniz viewed substantial forms as imparting a teleological structure to the appetites and perceptions of monads, structuring them according to lawful final causes. Adams (1994): 318.
[641] *R* 6:39-40.
[642] See Henry Allison's discussion of *Gesinnung* in Henry Allison (1990): 143.

because the infinite divisibility of space means that no fundamental and simple substance can exist in space.[643]

The idea that empirical laws are derivative entities that are shaped by an intelligible structure can be articulated in terms of an internalism, in light of which empirical laws are not true of the world in terms of correspondence with a God's-eye view on ultimate reality, but instead only within the spatiotemporal perspective of human sensibility.[644] These laws are *empirically* necessary, which means they are predicated on universal conditions of experience and not the ontology of the world in itself.[645] This distinction in metaphysical level of the empirical and intelligible levels can also be applied to address the two-selves objection we sometimes find in the commentary on Kant's metaphysics of agency, which also presupposes the idea that empirical laws and the free agent are in competition. The two-selves objection suggests that, in making an ontological division between the intelligible and empirical characters, we are separating the person into two distinct spheres, and making it incompatibly schizophrenic, insofar as it subject to empirical causation in its empirical character, but free at the noumenal level.[646] But, the two-selves objection actually assumes that the intelligible and empirical characters are ontological equals, pulling the self in opposing directions. The view I am

[643] We can only find an *analogy* of substance in space-time. Cited in Guyer (1987): 234.
[644] One can find an account of such internalism, which holds that truth is not correspondence to some absolute reality perceived by a transcendent knower, but instead is coherent within a rational system of beliefs, in Hilary Putnam (1998): 49-50. Internalism about laws is different from Allais's proposal that laws of nature hold only assuming a certain closed subsystem, e.g. billiard balls without people. Allais (2015): 306. Kant writes that "the understanding…is itself the legislation for nature…" *KdrV* A127.

[645] One can find an account of empirical necessity in The Postulates of Empirical Thought, *KdrV* A218/B265.
[646] Henry Allison criticizes the ontological division into a noumenal and empirical self in Henry Allison (2012): 87.

defending is, instead, that the lawfulness at the empirical level does not causally constrain the intelligible level, because the empirical level is internally realist within the human perspective. The intelligible level is the ground of this lawfulness, since it is an order determinate in itself, that is not relative to human sensibility.

*Diversified contingency involves neither causal inertness nor overdetermination***.** The intelligible ground that structures linkages in the empirical series is not causally inert, or a useless appendage that is epiphenomenal. Commentators have tried to incorporate anomalous monism[647] into Kant's theory of freedom, in which the intelligible level has a token-token identity with the empirical level and simply represents a perspective that abstracts from the lawfulness of the empirical level. If the intelligible structure is constitutionally identical with the empirical level, what is causally dominant is the empirical level, even though we may be able to re-describe the intelligible level apart from empirical laws. Instead of being a mere perspective, the intelligible character in my account is integrated as part of the systematicity of the world, serving as a foundation of the laws to which empirical events are subject. Perspectival views of practical agency, which leaves the possible sufficiency of the naturalistic world intact and do not acknowledge the possibility of a real intelligible ground, fail to integrate the moral significance of agency in a systematic way with the empirical order. Without qualifying the empirical level as metaphysically derivative, any agency from a perspective is a mere unsystematic fiction.[648]

[647] Donald Davidson (1980): 225.
[648] We find such a view of agency in Henry Allison, who thinks that there is no metaphysical fact of freedom, only a certain perspective. Henry Allison (2004): 48.

Henry Allison relativizes the naturalistic account of the person, in which it is controlled by its strongest desire, in virtue of a deliberative *standpoint* in which epistemic spontaneity of the mind allows the agent to conceive, though not experience, herself as taking certain inclinations as appropriate bases of action. The intelligible character for Allison is therefore mode of self-conception in the practical domain akin to the spontaneity of the theoretical domain in which the understanding subsumes individual impressions into tiered judgments.[649] But, the spontaneity here is not a metaphysical fact, but a mere perspective taken up in certain specialized contexts. Since Allison does not commit to a certain ontology of space and time, claiming that Kant offers instead an alternative to ontology (i.e. space and time are only epistemic conditions)[650], we cannot be sure the spatiotemporal domain does not exhaust reality. If freedom is a mere perspective incapable of metaphysical realization, it is not even clear how we can rationally take up this perspective. Freedom as mere perspective presupposes a fractured notion of theoretical truth, in which what is true is paradoxically relativized to a context. Merely because there are different justificatory strategies for freedom and the belief that one is determined[651]—i.e. that former is arrived through the practical perspective, and the latter through the theoretical—does not make belief in both rational, since holding something to be true seems to imply an independence from one's justificatory perspective. What is true, it seems, does not depend, by definition, on one's perspective.

An intelligible character that is a genuine metaphysically ground is not a causally inert perspective of dubious rationality. Nor is it the case that the intelligible ground that

[649] Allison (1991): 37.
[650] Allison (2004): 98.
[651] See Nelkin (2000): 574 for a discussion of the possibility of rationally holding contradictory beliefs.

forms maxims, in virtue of which certain linkages at the empirical level, between mental events and actions, are established, is a form of over-determination. Both causal inertness, and overdetermination, have the effect of making the intelligible ground of our empirical character dispensable. Overdetermination occurs when there is a causal closure of the empirical level, such that every empirical event has a sufficient empirical cause, and yet we posit another sufficient cause at the intelligible level. Two sufficient causes are not necessary, and so overdetermination leads to the rejection of the intelligible cause in favor of the observationally grounded empirical cause.

But, the metaphysics of agency I am developing does not involve the problem of overdetermination. It is true that an empirical event is the sufficient antecedent ground in the temporal series of action. Kant is not an indeterminist who posits gaps in the empirical series.[652] Still, the relationship of empirical events over time cannot fully account for itself. Empirical events require an inner ground, because the empirical appearance is a perspective-specific structure. The inner ground of maxims provides the structure in virtue of which certain empirical events become the sufficient cause of actions.

Advantages of the Metaphysics of TI in a Metaphysics of Agency

The metaphysics of transcendental idealism has several advantages over a flat metaphysics that absolutizes the empirical series of events, and denies any immaterialist

[652] *KdrV* A549/B577.

ground. Certain contemporary philosophers reject the idea of causation of events by substances, in the name of the scientific worldview. Substance-causality, according to David Velleman, would compromise "our scientific view of the world," which sees events as caused by other preceding events, or by nothing at all.[653] Universal event-causation is the thesis that all events are caused, and all causes are events—it is simply inconceivable that an event could be accounted for by a non-event.[654] The agent itself does not cause itself to act, but something the agent does, in the form of some temporally specific event. Introducing an agent-cause as a form of substance-causation without events forms an intolerable gap, it is claimed, in the scientific view of the world, that separates an agent's actions from the rest of the world, by leaving their arising unaccounted for by another event.

*Velleman's agent-equivalent mental state***.** David Velleman accordingly reframes agent-causation in the form of a mental state that is functionally equivalent to agency, without embracing any of the metaphysics of substance-causation. This mental state is a desire to act according to reasons, and so it is ontologically an event, although it has a special role of adjudicating between rational and irrational motivations that make it functionally similar to an agent.[655] Velleman's account may satisfy the thesis of universal event-causation, but it suffers from the same problems of contemporary libertarianism I documented in chapter 4: i.e. the agent's behavior is just the sum total of its psychological events, which are in turn the product of external events that cause them,

[653] David Velleman (1992): 467.
[654] Randolph Clarke also endorses the thesis of universal event-causation in Randolph Clarke (1993): 193.

[655] Ibid., 478.

and the agent lacks control of the interplay of these psychological events. The actions of Velleman's agent are not done *by* him, but done *to* him.[656] The agent lacks causal control of the interplay of psychological events filtered by this hypothesized agential mental state, and so Velleman's agent lacks *ultimacy,* an important term in the contemporary free will debate indicating an independent source-hood of action.[657] Identifying a mental state that promotes rationality imparts only a compatibilist form of control, since this state might exist even given determinism. We could incorporate alternative possibilities into this account, by making the rationality filter an indeterministic process, but these would be meaningless if the agent herself could not be the source of her action instead of a psychological process on which she is a spectator.

If we, through transcendental idealism, attach psychological events to an intelligible ground, structured by freely chosen maxims, then we can subject the interplay of event-causes to agential control. The series of events in the time-series has its structure in virtue of these freely chosen maxims, and so the agent is not a passive spectator on his own behavior.[658] We also preserve universal event-causation, since no event that characterizes the agent's empirical character lacks another event causing it. The only amendment is that we refrain from a dogmatic empiricism that makes events in time the ungrounded and exclusive features of reality.

Chisholm's dogmatic intuitionism. Roderick Chisolm relies on a pre-analytic faculty of intuition for establishing a certain ontology of the person, including its capacity

[656] Henry Allison's activity requirement on attribution of action to an agent requires that the action be done by the agent, and not done to it. Henry Allison (1990): 28.
[657] See Mele (2006): 13.
[658] Henry Allison (2004): 28.

to freely do otherwise than what it in fact does, and so bypasses the complex metaphysics of agent-causation in favor of a pre-philosophical mode of intuition. This mysterious pre-analytic faculty is not empiricist, since its commitment to a unitary personhood, which is more than a mere bundle of impressions, cannot be sustained by scrutiny of individual empirical impressions.[659] Nor is the pre-analytic faculty rationalistic, since it is just a baseline intuition rather than the product of argumentation, along the lines of Kant's proof of the ideality of space-time in the Antinomies.[660] In my account of the metaphysics of agency, one's practical cognition of oneself as absolutely free cannot be accepted as a self-evident proposition without at the same time uniting it systematically with an empirical order shown by theoretical reason to have a metaphysically dependent status in virtue of the antinomial conflicts.

Chisholm's account of agency is epistemically flimsy in comparison to the one developed in this chapter, since a self-presenting intuition of direct acquaintance with oneself, not based on rationalist or empiricist methodologies, can be dismissed by someone who either lacks the intuition, or who finds it defeasible with respect to an analysis of metaphysical structure.[661] Chisholm's claims to have direct insight into the substantiality of the self are vulnerable to skeptical attacks on the inference from the "I

[659] Hume showed that mere introspection cannot yield acquaintance with the self as an underlying substance—introspection, on the contrary, yields only acquaintance with passing mental events. David Hume (1949): I.iv.vi.

[660] Roderick Chisholm (1976): 23.

[661] Dieter Henrich identifies Kant's critical philosophy as a form of labor that ascends to the principle of freedom. The ascent in the critical philosophy is in contrast to the so-called mystics who descend from the principle of freedom after a mystical insight into its reality. I think Chisholm is invoking a mysterious and dogmatic faculty rather than taking the systematic approach of the critical philosophy. Dieter Henrich (2003): 67.

think" to the substantiality of the self Kant endorses in the Paralogisms.[662] Kantian humility recognizes that the mere fact of apperception—the mere fact that experience is united in self-consciousness—does not imply the existence of the self as a substance, let alone a self with absolute freedom.[663] What Chisholm claims is a noumenal knowledge of the self[664], insofar as direct acquaintance with the self means that there is no mediation by the forms of a particular sensibility. Chisholm conflates the awareness of the contents of the transcendental subject in inner sense—its passing thoughts—with knowledge of the substantiality of the self.[665] But, there is no more substance to this assertion, apart from a more rigorous justification, than an individual philosophical quirk, since other philosophers have very different intuitions, e.g. that events exhaust the activity of the self and the self lacks any substantial power. The Kantian can at least protect the inner nature of the self from the empiricist: the empiricist cannot dogmatically adopt a bundle theory of the self, or an event-causal one, because transcendental idealism means that individual empirical cognitions can never provide the full story of the nature of the self.

The lack of articulated metaphysics, evident in an un-integrated and dogmatic intuition of one's own agent-causal power, is a serious problem for Chisholm. Chisholm endorses a form of agent-causation in which, when the agent produces some new state affairs, the agent also contributes causally to its own production of the state of affairs. So, when an agent raises its arm, it not merely the case that this state of affairs is produced. If

[662] Karl Ameriks discusses how Kant's recognition of a stable unity of self-consciousness underlying different impressions is not enough to infer the substantiality of the self in Karl Ameriks (1981): 102.
[663] Kant shows this humility in the section on the Paralogisms, where he shows the fallacy in thinking that, because the "I" is the subject of all judgments, and never a predicate, it must be a substance. *KdrV* A349.
[664] Chisholm says that it is obvious we have direct knowledge of the self, and therefore we can bypass philosophical worries about the nature of the self. Chisholm (1976): 23.
[665] Chisholm (2013): 40.

action were the mere production of a state of affairs, it would be indistinguishable from reflex. When we deliberately will to raise our arms, we causally contribute to the production of this state of affairs.[666]

Even if one accepts this account of agency, the metaphysics that is supposed to undergird it are not clear. In virtue of what is the agent supposed to contribute causally to her own willing? We have a series of events: the raising of the arm, the agent making some exertion to raise her arm. But, it is not clear how this event-causal sequence is supposed to be attached to the agent. We could include another event, i.e. an exertion by the agent that causally contributes to her own exertion. But, we have the same problem involving an inability to attach the causal sequence to the agent, that is not solved by simply attaching antecedent events. If we claim that the series of events stop in some unconditioned event, we deal with the problem of an unconditioned event that may as well have arisen by chance, since it has no antecedent. We could continue to follow the causal series in an ongoing conditioned-condition relationship, only to end up with alien events external to the agent that show her to be merely a deterministic agent. If an agent causally contributes to its own production of behavior (since mere production could be a reflexive behavior and so not an action), does the agent also have to causally contribute to its causal contribution, to make the causal contribution an action, leading us into an infinite regress?

There is a philosophical bridge one has to cross, to move from events to the actual agent qua substance that is supposed to be the source of an event. Events will lead to a

[666] Roderick Chisholm (1976): 71. We can also understand the distinction here as between making something happen (i.e. mere production) and doing something (causally contribution to the production). See Emmett (1985): 76.

trilemma of unaccountable arising in an indeterministic process, deterministic event-causation, or an infinite regress of events in which the agent continually tries to gain control of its own undertakings.[667] The bridge from events to the agent is impossible to cross without the metaphysics of transcendental idealism.

TI, with its multi-faceted notion of the grounding relationship, can make crucial contributions towards saving agent-causation. The process that leads to the infinite regress relies exclusively on antecedent grounds. To account for how an agent causes event *e*, we posit an antecedent event, which is the event of the agent causing event *e*. However, instead of appealing to an antecedent event, we can introduce the idea of a constitutive ground, peculiar to TI, and so escape this unpalatable dilemma. TI permits the proposal that the agent causes event *e* not through some antecedent event, but by being the ground of it. Baumgarten mentions how the accidents of a substance require a ground, and that a substance grounds its accidents not as a prior event but through the power of the substance.[668] The agent, in the metaphysics of agency based on TI, serves as a structuring cause, in virtue of which a linkage is established between event *e* and another antecedent event—a desire or belief of some sort, or a combination of both. This constitutive ground solves the problem of control—i.e. it provides a metaphysical basis for the process of causal contribution, in virtue of which an agent takes ownership over

[667] One can find a discussion of this infinite regress problem in Chisholm's view in Timothy O'Connor (2000): 58.
[668] Baumgarten (2013): Sec. 197. Kant ties action and force to substance at *KdrV* A204/B249.

that to which it is causally contributing[669]—without introducing the problems related to events.

O'Connor's account of intrinsic action. Timothy O'Connor rejects Chisholm's idea that the agent has to exert a controlling event of causal contribution in order to be the cause of an event, claiming instead that agent-causation is distinctive in producing what are intrinsic actions. The mere exercise of an agent-cause is intrinsically an action, apart from some causal story connecting the event of the exercise to the agent.[670] For O'Connor, the agent's exercise of control does not itself require control to be considered an action.Even so, this formulation of agent-causation is useless if we cannot articulate a metaphysics whereby it is possible. Without the constitutive grounds provided by TI, agent-causation that is intrinsic, and does not involve a mediating causal contribution, reduces to the kind of intrinsic powers we find in non-causal libertarianism.[671] The non-causal libertarianism attributes action to volitional events that are simple, and so have no connection to the agent. If action is supposed to be intrinsic, it cannot be generated by some separate event of causal contribution, and so it simply arises, in a way that is detached from the agent. TI, on the contrary, allows the agent to directly own an action, without an antecedent event, in virtue of being the constitutive ground of the action—the

[669] "Taking ownership" is an idea similar to that of Frankfurt's conditions of moral responsibility, involving *endorsement* of one's will through second-order volitions. It is not enough that one's will causes one's action; one must endorse one's will to have ownership. See Pereboom (2001): 104 for a lucid explanation.
[670] O'Connor (2000): 59.
[671] We see the same problem in Griffiths, who claims that the notion of "trying," which is different from the actual action one does, gives insight into what is intrinsically under the agent's control. Griffiths argues that there does not have to be a special causal relationship between the agent and its own trying, because there is no need to try to try—the agent has immediate control over its own trying. But, there has to be some special way in which the agent is connected to its own trying, if the trying is not to be a mere chance event, along the lines of NCL, or a reflexive bodily movement. We can articulate this relationship through a Kantian metaphysics. Meghan Griffiths (2007): 24.

structuring cause[672] whereby certain antecedent events link up with certain actions. We cannot simply claim intrinsic agential control, in other words, without expanding beyond the thesis of universal event causation, so that we have another way of connecting the agent to its own action apart from events, i.e. as a constitutive ground. O'Connor claims that the event of agent-causation is different from a mere non-causal volition with a phenomenal quality of control, which we find in NCL, insofar as the former is a complex event (i.e. not a simple uncaused event) involving real causation of the agent and not based on mere phenomenality. However, some metaphysics of grounding contextualizing the exertion of the agent-cause is needed to secure unity of the person, i.e. prevent the agent-cause from itself being an uncaused event. Otherwise, even the complex event of agent-causing-x will itself lack grounds and act as a spontaneous event detached from the agent.[673]

Kant also critiques the thesis of universal event causation, in order to provide an account in his *Religion* of how we might be responsible for the character in light of which we make individual choices. Human nature for Kant is not an antecedent ground in time, some basic natural impulse that pushes the agent towards selfish motivation. If human nature were such an event—some built-in drive towards concupiscence that makes us like what is bad for us—then we could not be responsible for our own selfish tendencies.[674] Rather than being some antecedent ground, the natural tendency towards selfishness that characterizes human beings is a subjective choice that underwrites every deed of ours that

[672] We find the idea of a structuring cause in Fred Dretske, who uses the example of wiring in a doorbell system that permits the linkage of the pressing of a button and the ringing of the doorbell. Structuring causes establish a causal pathway between two events. Fred Dretske (1988): Chapter 2.
[673] It will not be a "personal fact," to use a phrase from Watkins (2005): 303.
[674] *R* 6:21.

falls in the empirical time-series. Without a sophisticated metaphysics, it is impossible to effectively capture how an agent is supposed to define itself freely on a fundamental level, in terms of a general character that imparts a tendency towards selfishness in all of its actions. Using only antecedent grounds, we would have to identify this fundamental subjective maxim either with an innate impulse with which an individual was born, for which the individual could not be held responsible. Or, we could identify this maxim with a fatalistic choice an agent makes early in life, from which all its subsequent actions flow—a wildly implausible hypothesis. Instead of situating it in time, TI allows us to establish the subjective umbrella maxim at the constitutive root of the empirical series, so that, for every action in time, there is a choice of character that underwrites it.

Objections to Agent-Causation in Light of Metaphysical Transcendental Idealism

Objection 1: Intelligible grounding relationships represent an overextension of the categories beyond possible experience. Any metaphysical grounding relationship between the intelligible and empirical selves has to deal with the epistemological problems associated with *noumenal affection*. Noumenal affection is the idea that the content of experience is generated not by empirical objects in space and time, but by unknowable, non-spatiotemporal things in themselves.[675] It seems as though TI requires noumenal affection, since TI designates the spatiotemporal world as a mere appearance and as demarcating the limits of our cognition. Accordingly, the spatiotemporal world must have an unknowable and non-spatiotemporal ground.[676] On the other hand, TI

[675] Desmond Hogan (2009): 502.
[676] Kant frequently states that appearances must have grounds. *KdrV* Bxxvif.

seems to prohibit noumenal affection, since TI seems to prohibit any trans-phenomenal use of the categories for cognition.

However, this objection is uncharitable. First, it is incorrect to restrict both mere speculation, and cognition, through the categories, exclusively to appearances. Cognition requires an intuition of objects, and we can have intuitions only as mediated through the spatio-temporal structure of our sensibility. While cognition is bound to our sensibility, thinking has an unbounded field.[677] It is not incoherent to claim that we cannot cognize things in themselves, because they can never be given directly to our sensibility, but we can think of them. Denying even the thinkable intelligibility of things in themselves fails to take into account Kant's endorsement of the coherency of the intuitive intellect. The intuitive intellect is able to have direct intuition of things in themselves through its understanding alone, unmediated by sensibility. It is a coherent possibility against which Kant compares the human discursive intellect.[678]

There is a danger in deploying the principle of significance so as to destroy metaphysics, of conflating meaning with justification, or cognition. That is, it is incorrect to argue that, because we cannot have cognition of what is outside of space and time, pure categories divorced from sensibility cannot have meaning. The simple fact that pure categories are distinct from one another—e.g. substance, quantity, and quality are clearly distinct from another—and that they are based on logical functions that are defined by a complete detachment from any intuition at all, shows that they have meaning apart from

[677] *KdrV* B166.
[678] The intuitive intellect is discussed in *KdU*, Sections 76 and 77. Guyer discusses how Kant builds his epistemology against the background of the theology of a creative intellect that has absolute knowledge of its creations in Guyer (1987): 5.

intuition. There needs to be a Transcendental Deduction because a priori concepts, which have their meaning independently of experience, may or may not underwrite empirical intuitions.[679]

The sort of immaterialist substructure to the event-causal order in our agency is not proven, in the manner of dogmatic rationalism, through cognition of theoretical reason. Instead, theoretical reason's task is merely to relativize its own empirical cognitions, in light of the Antinomies generated when reason's innate grasping for the unconditioned conflicts with the categorical structure of the understanding. It is practical reason that is then able to extend theoretical cognition, in a systemic manner that is not a mere privately valid whim, into the supersensible capacity of absolute freedom.[680]

The notion of a thing in itself comes from a fundamental stance of interactionism between mind and world, which Kant preferred over the more baroque causal theories of pre-established harmony, in which interaction between mind and world is replaced by a parallelism with divinely ordained harmony, or an occasionalism in which God is the sole cause and finite substances lack any causal interaction intrinsic to them.[681] The basic metaphysical belief in mind-world interaction, involving a causal relationship between the mind and some finite substance outside the mind, is more fundamental than the spatiotemporal framework of our experience. Since we are acted on by an independent finite substance, and this generates our experience[682], it makes sense that the

---

[679] Karl Ameriks effectively rebuts the error that concepts cannot have meaning apart from intuition in Karl Ameriks (2003): 28.
[680] Kant discusses practical reason's access to the supersensible in *KdpV* 5:4.
[681] For a discussion of these different theories of causation, see Karl Ameriks (2003): 25.
[682] In R5653, Kant claims that representation in space requires determination of the mind by an ontologically independent thing. Cited in Guyer(1987): 318. In this inference (not a mere default starting-point), Kant appears to be overstepping the boundaries of his Critical epistemological restrictions.

metaphysical origin of our experience is more fundamental even than the most basic structures of our experience. The fundamental belief in mind-world interaction ultimately leads to a common philosophical ground, shared in typical discourse, that the framework of our experience is finite and that it cannot therefore embrace in their entirety the substances that generate our experience. This philosophical common ground is not a dogmatic inference from our experience, which would involve an illicit use of the Second Analogy to track the conditions of our experience into the supersensible. Interactionism between the mind and finite substances, and the subsequent acceptance of the limitations of the framework of our experience, is not a dogmatic inference but a recognition of a common default position that is more natural than the esoteric metaphysics of pre-established harmony or occasionalism.[683]

The idea that there is metaphysical interaction between the mind and the world is perhaps a basic belief that reaches to the very limits of the framework of discourse within which philosophical speculation takes place. One can rationally entertain a Leibnizian position of pre-established harmony, or an occasionalist theory of causation, but such theories can only be entertained in philosophical speculation, and one always reverts in practice to a baseline conviction of mind-world interaction. The idea of mind-world interaction is the default position from which someone has to be argued out of—there is no one who starts out as someone who believes in pre-established harmony, or occasionalism, and then has to be convinced that finite substances interact with one

[683] Ibid., 26.

another. Mind-world interaction forms the very framework within which we understand counterintuitive theories about causation like pre-established harmony and occasionalism.

The belief in metaphysical interaction, then, is a sort of anchoring belief that is a constant presupposition. Metaphysical interaction might not be grounded in an intellectual way that is foolproof against skeptical doubts based on speculative alternative theories, but it surely is grounded in a way of life.[684] Kant is therefore absolved of unduly dabbling in dogmatic metaphysics that transcend the limits of possible experience in positing the thought of things in themselves as absolute grounds of experience. The belief in metaphysical interaction is not a dogmatic inference but a point of orientation based on a recognition of a philosophical starting point of ordinary discourse.[685]

Agent-causation is incompatible with the scientific worldview. This historical debate over dogmatism in the thesis of metaphysical interactionism manifests itself in the contemporary discomfort with agent-causation we see in theorists like Kane and Balaguer. Just as the historical debate is fearful of an extension of the pure understanding in a merely speculative manner unconstrained by possible experience, contemporary objectors to agent-causation fear the empirically unjustified assertion of special capacity for an active power, capable of self-determination in ways that betray the lawfulness of the empirical order and transcends ordinary event-causation. Kane observes the principle that we should not incorporate into our metaphysics anything beyond that to which a

[684] This argument about how a belief in interaction is a fundamental anchor of a way of life, such that, though one might skeptically probe it and introduce alternative theories of causation, it persists as ineradicable, is partially inspired by Ludwig Wittgenstein's work *On Certainty*. Ludwig Wittgenstein (1969): see especially sections 113-115.

[685] Ameriks (2003): 23. Kant provides an argument in his pre-Critical period for mind-world interaction: the substance of the mind cannot undergo changes in isolation. *Novo* 2:411-412.

compatibilist theory of free will would appeal. A compatibilist theory of free will needs only appeal to temporal events passing through one's psychology[686], and so it does not need to take on the metaphysics of interactionism operating prior to spatiotemporal intuition. The idea here is that it would be awkward to approach the free will problem with a methodology that respects empirical discoveries emerging from science, while at the same time positing, in one's final theory of free will, what is, to science, the very strange entity of a non-empirical agent-cause. Contemporary theories of free will seem to take a Humean attitude to metaphysics: any non-empirical metaphysics of agency involves a dabbling in empty speculation that needs to be disciplined by a strict adherence to simple impressions in experience.[687] Kant notes that sensations alone occupy instants[688], and so a Humean sensations-based epistemology cannot articulate the substance-metaphysics associated with character that involves a unity of person across time.

The problem, though, is that, as I showed in Chapter 4, severing empirical data from a larger context of metaphysical interaction ends up with a metaphysics of agency that can only merit the objective attitude. The decisions of the ECL agent happen to it, and intellectual honesty requires us to interact with the ECL agent in the same way as we would interact with a psychiatric patient, or a malfunctioning machine. An agent cannot be truly involved in decisions that are the simple fallout of indeterministic processes, and

[686] Compatibilist theories of freedom only require freedom of action (i.e. absence of impediments), rather than some control over the sufficient reason of one's action and some ability to do otherwise. Derk Pereboom and Michael McKenna (2016): 50.
[687] Dieter Henrich mentions how the permeation of Hume in German culture made metaphysical questions into pseudo-problems that an empiricist commitment to simple impressions could safely dismiss. Henrich (2003): 27.
[688] *KdrV* A167/B209.

interacting with such an agent would involve manipulating the agent so as to redirect behavior, rather than negotiating with the agent in order change behavior which is not, as an indeterministic process, amenable to change overseen by the agent.

The principle of significance is the principle that attributes significance to pure concepts only when they are wedded to some empirical content. We can use this principle to atomize experience, and sever individual empirical cognitions from an underlying metaphysics of causal interaction between substances responsible for generating experience, and not fully expressible in spatiotemporal empirical terms.[689] However, one of the conclusions of this dissertation is that this sort of aggressive empiricism, which dogmatically rejects even the mere thought of non-empirical metaphysical structures, takes empirical data out of a context essential to them, and actually discombobulates the cognitive project in the name of making it empirically grounded. We have seen how, unless space-time is taken as ideal, any attempt to articulate empirical cognition in terms of conditioned-condition relationship in accordance with the principle of sufficient reason leads to the contradictions of the Antinomies. The series of empirical events in the time-series leads backwards according to the rational demand for conditions, only to find that reason's demand for completeness conflicts with the demand that each individual contingent condition have its own condition. We see, too, that any sort of structure in empirical science requires non-empirical elements of cognition, by which individual empirical cognitions are subsumed into the stable and lawful structures, e.g., the

[689] Strawson sees such a contradiction between the epistemology involved in the principle of significance, and the larger metaphysics—which Strawson rejects—involving a subject-dependent experience peculiar to human nature set against the backdrop of a fundamental reality of things in themselves that are not dependent on human sensibility. Peter Strawson (1966): 5-9.

Analogies of Experience.[690] We know from the *Metaphysical Foundations* that Kant rejects the idea that the mere existence of matter can account for why it fills space and is impenetrable, opting instead for a dynamical theory of matter which accounts for impenetrability and the occupancy of space in terms of more fundamental repulsive and attractive forces.[691] Of course, the specter of empirical chaos in the organization of empirical cognitions into lawful structures is not solved by the transcendental structures of the understanding, which only impart a structure to nature in general, but in virtue of the faculty of judgment, which has the special task of mediating between the theoretical cognition of the understanding and practical reason's grasp of the unconditioned, in the form of its absolute freedom.[692] Again, spatiotemporal reality is shown to be metaphysically subordinate, insofar as the very unity of our experience in spatiotemporal reality depends on the coherent integration of it into reason's orientation towards the supersensible. The empiricist attempt to divorce spatiotemporal fragments from their context actually eviscerates the context that makes them intelligible.

*The problem of timeless agency***.** The picture of metaphysical interaction, in which a relationship of causal interaction between the mind and external finite substances generates experience and so is more fundamental than the spatiotemporal framework of experience, generates advantages for a metaphysics of agency. The notion of an intelligible ground provides depth to an account of action, in virtue of which we can connect the agent to the sequence of events transpiring in its psychology, making it the

[690] They begin at *KdrV* A176/B218. These are presaged by the Principles of Harmony in *Diss* 2:418.
[691] Kant is criticizing the view of Lambert, whereby impenetrable occupancy of space is an inherent feature of empirical objects. Immanuel Kant *MadN* 498.
[692] *KdU* 5:196.

overseer, in the form of an underlying character, of the prioritization of certain reasons over others in patterns connecting beliefs-desires and actions. In addition to avoiding the problem of a disappeared agent, by supplementing the sequence of events that merely happen to the agent, the fact that the agent's intelligible character is outside of time gives it a freedom to act otherwise that empirical contingency in time, which is subject to an antecedent event in the past, does not have. Removing the intelligible character from the empirical series, so that it can act independently and not just heteronomously respond to external empirical stimuli, is essential to Kant for grounding genuine virtue. Genuine virtue cannot merely be a habit one falls into because of environmental influences, or something one does as a means to some empirically determined end. Virtue has to be developed through moral judgment that is independent from empirical causality.[693] This moral judgment has to be metaphysically efficacious, moreover, and cannot be merely epistemic, since any action of a merely epistemologically spontaneous agent will be determined heteronomously by empirical causes that metaphysically dominate the deliberative perspective.

However, this diversification of the contingency relationship leads to the perhaps more serious problem of causal transactions occurring between a non-spatiotemporal base and an empirical domain in space-time. Causal interaction between space-time and what is not in space-time seems paradoxical, insofar as causal interaction with the spatiotemporal field would seem to require some sort of spatiotemporal location. This tension is especially poignant with respect to agency. It seems impossible for an agent to incorporate certain sensible inclinations into its maxims, if the intelligible character of the

[693] This issue is discussed in Henry Allison (1990): 165.

agent is supposed to be outside of time, whereas sensible inclinations occur at specific times.[694] Nevertheless, Kant writes, "Pure reason, as a merely intelligible faculty, is not subject to the form of time."[695]

If the timeless agent is not able to intervene in individual situations, responding to the arising of sensible inclinations, then it seems as though we have to develop an account of agency in which the agent makes a timeless choice that amounts to a choice for the entire empirical history of the agent. Kant states that the intelligible character exists only as a mode of thought (*Denkungsart*) and so is distinct from the spatiotemporality characteristic of the mode of sense pertinent to the empirical character (*Sninesart*). As something abstracted from spatiotemporality, nothing arises or begins to be in the intelligible character.[696] Allen Wood accordingly portrays Kantian free will as a single timeless choice that is simultaneous with every single choice the individual makes over the course of her life. Since this choice is of an entire empirical history, it raises the problem of fatalism: every choice one makes on the empirical level is the unfolding of this single timeless choice, and so one is on a pre-determined course.[697] Moreover, the idea of a single simultaneous cause of one's moral history involves wild revisions of our common-sense view of agency.

But, Wood's account of the timeless agent, though it at least proves that timeless agency is logically possible, dogmatically restricts the Kantian's options for articulating the relationship between intelligible and empirical levels. A more effective way to

---

[694] Henry Allison makes this objection in Allison (1990): 52. See also Jonathan Bennett (1974): 226. Hud Hudson makes a similar point in Hud Hudson (1994): 26.
[695] *KdrV* A551/B579.
[696] *KdrV* A551-2/B579-80.
[697] Allen Wood (1984): 73-101.

incorporate this intelligible character into an empirical account of action in time is not to appeal to a single simultaneous cause, but to appeal instead to some possible noumenal analog of time in the empirical series by which the intelligible character could exercise multiple choices. Wood's timeless agent exercises only a single overarching choice because simultaneity with every empirical act requires a single choice. But, it is questionable whether the timelessness of the noumenal must manifest itself as simultaneity.

The thought that either the temporality that characterizes the empirical series characterizes the noumenal level, or no time at all characterizes it, betrays a dogmatism that seeks to limit the nature of objective reality in itself to the subjective principles of our empirical mode of cognition. A true disavowal of dogmatism would recognize that there might be a form of temporality that characterizes the noumenal domain that is different from what characterizes the empirical domain.[698] Though we cannot have cognition apart from sensibility, we can think about non-sensible reality through the categories. We know enough, from various parts of the Transcendental Analytic, about the structure of pure thought to articulate a theory of a possible analog of temporality at the intelligible level, which would not lead to the absurd metaphysics of a single simultaneous choice of the entire course of one's life. Kant notes that we visualize time as a spatially extended line, and it seems as though we could distinguish between this specialized version of time, involving piecemeal progression, conformed to human sensibility from an abstract noumenal time detached from human sensibility.[699]

[698] Kenneth Westphal makes this point in Kenneth Westphal (1997): 231.
[699] *KdrV* B154.

Kant distinguishes between general logic and transcendental logic, in virtue of the fact that the latter undergirds the a priori conditions of sensibility, space and time, whereas the former abstracts from all content.[700] So, in transcendental logic, in which we are dealing with the concepts meant to underly objects of sensible intuition, we have a concept of cause. In general logic, corresponding to the concept of cause, we have a logical function, applicable beyond the spatiotemporal framework, of a hypothetical judgment.[701] The hypothetical judgment is an asymmetrical grounding relationship that is abstract, insofar as it is not transposed into the empirical time-series.[702]

The logical structure of ground-consequence can provide the underpinnings of an account of a temporal analog at the intelligible level. The intelligible character, Kant tells us, has causality, although it is intellectual causality that does not stand under empirical conditions.[703] The intelligible level does not have to cause an empirical act as an antecedent ground in time, or as a simultaneous one. Instead, the intelligible character can act as an underlying character that serves as a hierarchal formal structure (i.e. not efficient causes in time) constraining the empirical agent to act in terms of the basic values of its character. Abstract ground-consequence relationships, in the form of hypothetical judgments of general logic abstracted from time, can establish a sequence in virtue of their asymmetry. The fact that there is a sequence at the intelligible level opens up some flexibility in virtue of which we could correlate the temporal sequence at the

[700] *KdrV* A77/B102.
[701] *KdrV* A70/B95. See Henry Allison (2004): 224.
[702] *KdrV* A74/B99, A55/B79.
[703] *KdrV* A540/B568.

empirical level with that of the intelligible level. Sequence characterizing intellectual causality would allow us to avoid Wood's simultaneous meta-choice.

Wood's idea that the intelligible character makes a single choice that is simultaneous with every empirical action actually smuggles in empirical temporality into the intelligible level, and so is guilty of the dogmatism about temporality mentioned above.[704] In Inaugural Dissertation, Kant recognizes that simultaneity is itself a feature of empirical temporality. Simultaneity is a perpendicular feature of time in the phenomenal world, that shows things existing in the same time-slice, to complement the linear succession of time through different time-slices.[705] Simultaneity is also one of the three modi of time in empirical cognition for which the Analogies are supposed to establish objective grounds.[706] A noumenal analog to time can abstract from the modi of time peculiar to human experience, i.e. both succession and simultaneity.

It is better to articulate the sequence at the intelligible level in terms of an abstract ground-consequence relationship from which both succession and simultaneity have been removed. The choice of the intelligible character, on my account, is not some single founding choice constraining the entire history of the empirical character, but instead a ground of the empirical character that constrains the kind of events that flow through it, while at the same time transcending the temporal determinations of the human observer. Since there can be a noumenal analog to empirical time, the noumenal choice does not have to be a single immutable choice simultaneous with an entire empirical history. Instead, there can be a series of noumenal choices in a sequence that is neither

[704] Wood (1984): 96.
[705] *Diss.* 2:401.
[706] *KdrV* A177/B220.

simultaneous nor successive. A series of noumenal choices is more plausible than a single meta-choice, and a noumenal analog to time, that does not conflate timelessness with simultaneity (which is actually an empirical determination), can provide the basis of this series of noumenal choices.

The abstract ground-consequence relationship can assure that, as Kant puts it, nothing happens at the intelligible level, while at the same time this abstract relationship of the pure logical structure of thought in general can sustain a sequence. The maxims at the intelligible level in virtue of which an individual shapes her empirical character can all exist, so that nothing happens at the intelligible level in the sense that nothing comes into existence in time. Nevertheless, the fact that nothing happens at the intelligible level is compatible with things existing in an asymmetrical sequence. A God's eye view on the series of maxims chosen by an individual in relation to her intelligible character would show the entire distribution of them at once, since God is not limited to the piecemeal progression of time associated with human sensibility (i.e. the linear notion of time). At the same time, the God's-eye view would still be able to identify a sequence in relation to the maxims, structured according to asymmetrical hypothetical judgments of general logic.

Kant talks about a revolution in one's intelligible character, such that an individual comes to respect duty for its own sake. In the revolution, the foundation of one's maxims shifts from an impurity in which the moral law is used as a means to achieve some empirical incentive, to a pure foundation in which respect for the moral law

is sufficiently motivating.[707] One might think that, in order for an intelligible conversion to take place that also revolutionizes one's empirical character, the revolution at the intelligible level would have to arise in time in correspondence with the change in the empirical character, as a simultaneous accompaniment. However, we parse the empirical/intelligible distinction in terms of a distinction between intuitions that arise in time and an intuitive intellect that bypasses the sensibility and so is not restricted to the temporality. By extension, we can parse the distinction, between the intelligible change in maxim and its empirical manifestation, in terms of maxims that always exist in a fixed sequence and manifestations at the empirical level that arise and perish in the passage of time through the present moment of an individual observer. Kant, in fact, hints at such a distinction by recognizing that a choice of a new maxim, that reverses an impure privileging of the empirical towards a pure respect for the moral law, is recognized by God at the level of the intelligible ground of behavior, but only manifests itself empirically in a gradual progression towards holiness. God is able to bypass the conditions of human sensibility, and see directly into the intelligible ground of an individual's maxims.[708] The choice of duty over selfishness comes after, then, the choice of a maxim privileging selfishness, and they are discrete choices rather than a single overarching choice. But they exist in an abstract sequence that is different from the arising in time of gradually improved forms of behavior at the empirical level.

The idea that practical agency involving the incorporation of empirical impulses into maxims is incompatible with timelessness actually comes from a monolithic

[707] Kant here incorporates the New Testament language of a "new man." *R* 6:47.
[708] Ibid., 6:48.

understanding of grounding relationship from which Kant broke free with the Critical turn. The objection to timeless agency presupposes that incorporation of an empirical motive has to occur as an antecedent ground in time. The act of incorporation is just another step in the desire-belief sequence leading up to action. Alternately, along with Wood, incorporation has to occur through a mysterious timeless choice that is occurs *at the same time* as each empirical action. But one of the upshots of transcendental idealism, which we see especially in the Antinomies, is an opening up of a supersensible space by designating the series characterized by temporality as merely ideal. The supersensible serves as a different kind of ground than the antecedent grounds in time—it is not one that precedes or coincides with the empirical series in time, but instead is a sustainer of the being of the empirical series at an absolute level detached from the temporality of human sensibility.

A fundamental asymmetry in grounding, in virtue of which a maxim at the intelligible level grounds a sequence of events at the empirical level and not vice versa, accounts for the direction of time, in which certain past mental states are the antecedent grounds of future actions.[709] The actual sequence of events leading up to someone's action are unaccounted for unless there are causal powers constraining the sequence to actualize certain possibilities and not others. Causal powers are defined in terms of possibilities which they exclude. A merely actualist metaphysics, which ascribes reality only to actually occurring events, cannot account for why certain events occur and not others. Actualities that are unaccounted for represent a fragmentary world in which

[709] Gregg Rosenberg develops an account of how the direction of time is dependent on asymmetrical causal relationships—themselves not in time and ontologically prior to it—bringing potentiality into actuality. Gregg Rosenberg, (2004): 214.

contingent events just arise in a free-floating manner, and has deleterious implications for the unity of the agent.[710]

*Introducing concepts from the philosophy of time***.** The Kantian metaphysics of agency involves actual events in space-time, and accounts for them in terms of maxims at a non-spatiotemporal intelligible level that, in virtue of causal powers in the form of abstract ground-consequence relationships, constrain certain movements from indeterminate possibility to actuality. We can use terminology from the contemporary philosophy of time to illuminate this metaphysics. The fact that events in the empirical series both arise and perish as the agent progresses along the linear path of time indicates a presentism[711] about the empirical series, in virtue of which only the present moment exists. The fact that the empirical series is merely ideal—i.e. a mere appearance—means it can only exist in the present moment of empirical synthesis for a human subject. Kant says that appearances cannot be given in themselves, but can only exist as perceptions.[712] The empirical series, moreover, cannot exist as a whole, even if we were able to view the whole extent of it from a God's-eye perspective. This is because there can be no unconditioned on the empirical level, since, on the empirical level, events are always subject to an antecedent ground in time.[713]

We should accordingly supplement the empirical series, which consists only in actual events occurring in the present moment[714], with an underlying structure in which

---

[710] Rosenberg argues that causal powers require an abstract modal realism about possibility, because if there are no possibilities that a causal power is constraining, it is causally idle. Gregg Rosenberg (2004): 207

[711] A definition of presentism can be found in Kristie Miller (2013): 346.

[712] *KdrV* A494/B522.

[713] *KdrV* A483/B511.

[714] On the close relationship between actualism and presentism, see Theodor Sider (1999): 326.

the unconditioned might exist and which therefore might represent a genuine whole. Eternalism is the view that not only the present, but also the past and future exist as well. Eternalism thus incorporates into an actualism about present events, a metaphysics in virtue of which events in the past or future can still exist.[715]

We can represent the intelligible structure in terms of eternalism. The intelligible structure cannot be presentist, because it is not tied to down to a particular sensibility in virtue of which only the present moment of empirical synthesis exists. Though eternalism is detached from any observer, it nevertheless maintains a sequence, so that there can still be an asymmetry in the intelligible structure by which some parts come before, and some parts come after, one another. The sequential relationships at the intelligible level would be permanent and objective, as opposed to the impermanence and subjectivity of the present moment in presentism. Whereas, for the presentist, the empirical level involves regular motion through the present moment, so that future flows into the present moment (it arises), and the present moment flows into the past (it perishes), the sequential relationships at the intelligible level are permanent. For the presentist, what is past, present, and future varies according to the subject—what is past for me is future for someone else. For the eternalist, before and after are objective relationships.

So, Kantian metaphysics of agency involves both an anthropocentric perspective, in which things arise and perish with respect to a human observer, and a theocentric perspective, in which there is a stable sequence of before and after. Nothing happens or arises in the theocentric series, since it all exists at once: it is not as though events move

[715] Ibid.

into and out of the present moment for the divine being, because the divine being has no present moment.[716]

My critique of Wood, which I began in the previous section, can here be completed, in light of the distinction between the anthropocentric and theocentric perspectives on time. For Wood, the timeless intelligible character chooses an empirical history from a set of possible histories. In order to avoid a fatalism, in which an agent's single choice of a history makes her actions an inevitable product of the pre-selected past, Wood posits that the timelessness of the choice of history is simultaneous with each individual act. Wood's choice is a single unitary choice from a God's-eye point of view in light of which each empirical act in time is immediately present. Wood's synthesis of a timeless choice and a deterministic empirical series is thus a form of altered-past compatibilism, in which free agency exercises alternative possibilities by selecting a different past. Selecting a different past does not violate a deterministic causal sequence—actions still flow out deterministically based on the choice of one's empirical history—and so this is a form of compatibilism.

Wood's view, though, smuggles in temporal determinations that properly belongs to the anthropocentric series. Positing the timeless choice as a single once-and-for-all choice *simultaneous* with all empirical acts situates the timeless choice into the phenomenal world of empirical time. The implausible single unitary choice is forced on us because of the conflation of simultaneity and timelessness. Multiple choices, Wood seems to presuppose, would require succession. The view I present in this dissertation is a

[716] We can find another way to articulate the difference between objective eternalist structures and subjective presentist structures in the distinctions between the different series of time in J. Ellis McTaggart (1908): 459-60.

form of altered-law compatibilism.[717] Here, free will does not interfere with determinism insofar as it only changes laws, and does not generate events that violate those laws. Altered-law compatibilism does not frame noumenal agency in terms of the anthropocentric series, as *simultaneous*, one of the modi of schematized temporal cognition, with each empirical event. It is the attribution of simultaneity to the noumenal agent that forces the idea of an altered past, i.e. a single meta-choice. Instead, the free agency in altered-law is the supersensible ground of the empirical character beyond schematized temporality, in virtue of which certain laws link certain event-types at the empirical level. Free agency does not figure in the present moment, because it is not relative to an observer with a certain temporally conditioned sensibility. Even so, the free agent in altered-law compatibilism can exercise its causality in multiple acts in a sequence, insofar as exemption from the past-present-future sequence relative to a human observer does not preclude an abstract non-unitary sequence.

Rather than being a causally inert form of agency, relegated by its timelessness to an irrelevance from the give and take of empirical inclinations, the eternalist objective structure can actually link up with the empirical character in virtue of a new form of grounding introduced by transcendental idealism. A restricted view requires that the only way to account for action is through an antecedent ground in time, which either precedes or is simultaneous with its consequence.[718] Allison presupposes this restricted view in thinking that only way to incorporate inclinations into maxims is by exercising causality

[717] Hudson discusses the difference between altered-past and altered-law conceptions of freedom at Hudson (1994): 33.

[718] Kant identifies an *identical* ground in *ND* as a kind of antecedent ground, which grounds a predicate by being identical with it. E.g. a triangle is the identical ground of its through sides. Antecedent grounds, which account for the existence of a consequence, can both precede and be simultaneous with their effects. *ND* 1:392.

in time. Instead, transcendental idealism, as I have noted, demotes the empirical series in space-time to mere ideality, and restricts it from being a whole. The objective eternalist structure, detached from the specifically located human observer, represents the ground of the empirical series. Nothing happens or arises in it, since it is detached from a present moment and the concomitant non-actualized past and future. Empirical time is not an absolute reality but instead a bridge which, in its a priori universality, connects the pure categories to empirical intuition.[719] In spite of the lack of arising at the eternalist level, there can still be an asymmetry such that one thing comes before another. These asymmetrical structures, expressed in abstract hypothetical judgments, form the basis for the sequence at the empirical level. One event on the empirical level is in the future, while another is present, and yet another is past, in a way that is grounded on certain objective asymmetrical structures that do not have a relationship to a specifically located subject. These objective asymmetrical structures are not limited to single choices that ground everything all at once, because we are not importing the empirical time-determination of simultaneity into the noumenal level.

One might understand this metaphysics in terms of a God's eye perspective, from which God knows all of one's actions even before they occur in the present moment in which one can observe them. Nothing arises in the God's eye view: one's action when one is ninety is available to the intuitive intellect, just as one's action as a three year old is. At the same time, one has a discursive intellect in which experience is mediated through a sensibility with a spatiotemporal structure. One does not get the God's eye

[719] Time is the schema of the concepts of the understanding. *KdrV* A138/B177.

view on the full extent of one's actions, but only those that happen to exist in the present moment of active perception.

The empirical series is controlled by antecedent grounds in the past, and one lacks control over grounds that are in the past—one can only act in the present. Structures in the intelligible level are not subject to antecedent grounds in the past, because the intelligible level is not in the empirical time-series. The individual's free acts at the intelligible level are contingent in the sense that they could have been otherwise, but not in the sense that they require some antecedent ground in the past. The agent has a special sort of control over its intelligible character because it is not controlled by a past in which it can no longer act.

*The problem of contrastive explanations***.** A charge in contemporary debates against agent-causal theories is they are unable to ground *contrastive explanations*. A contrastive explanation explains why an event occurs at one time and not at another time. Agent-causation is unable to appeal to a temporally located state, because its central claim is that the agent causes action, and not an event.[720]

However, my account of agent-causation has resources to address this problem. Again, the key to the solution is in diversifying our understanding of grounds. At the empirical level, the action of the agent *does* have an antecedently determining ground in virtue of which an event at a certain time causes an event to arise. The empirical world is governed by the Second Analogy, in virtue of which objective succession requires a causal law.[721] Accordingly, the appearance of an action is preceded by an empirical event

[720] Carl Ginet: (1997): 93-4.
[721] Temporal succession is the schema of the pure form of causality. Guyer (1979): 374.

that is the basis for a contrastive explanation—e.g. I went to the store *today* because *this morning* I saw there was no bread. In addition, though, the antecedently determining ground at a specific time cannot fully account for the action, because it also requires a constitutive ground in the form of a basis that is dis-affiliated from the human perspective. The diversification of the notion of ground allows for contrastive explanations without at the same time reducing our explanation of action to mere events that, whether they are deterministic or indeterministic, make the agent a mere spectator.[722] Diversification of grounding can, in fact, address objections to free will based on Libet experiments: the presence of an unconscious readiness potential prior to action does not preclude the agent from serving as the intelligible ground of the linkage between the readiness potential and action.[723]

In *KdrV*, Kant calls the empirical character the *sensible sign* of the intelligible character.[724] The intelligible character does not figure in the empirical series as yet another antecedent ground in time. The intelligible character represents the constitutive ground of what appears as an empirical character only from an anthropocentric perspective.

The agent qua substance acts as a temporally indeterminate ground, which provides a structure, through maxims, to empirical events. The maxims serve as umbrella maxims that govern a host of actions in virtue of their temporal indeterminacy. Henry

---

[722] Constitutive/intelligible grounds mean that the empirical discovery of determinism, contrary to what van Inwagen says, would not force us to adopt compatibilism. Deterministic causation is intact on my account, but not complete. Cited in Mele (2006): 163.

[723] The Libet experiments are discussed in Mele (2006): 33. The idea here is that an unconscious readiness potential prior to a self-aware choice means that actions are not freely chosen—the brain chooses them for us.

[724] *KdrV* A545/B573.

Allison notes that maxims in relation to practical reason perform a role parallel to what empirical concepts play in relation to sensible intuition.[725] The same empirical concept can unite various sensible manifolds across time. Similarly, maxims unite certain beliefs and desires, in light of a normative principle, with certain actions across time. Maxims, moreover, provide a texture to the agent-substance that prevent the mere existence of the substance itself from guiding its empirical character.[726] Instead, maxims represent grounds within the agent-substance. This texture further distances us from a crude substance-causation model that simple ties an event to a temporally indeterminate and untextured substance. The grounds of a substance are neither the mere existence of a substance, nor are they themselves events. Instead, grounds/maxims are an activity of an agent-substance by which it grounds the flow of empirical events.[727]

The objection that agent-causation causes a temporal event without itself having any temporal specificity is related to another objection we find, namely, that, because agent-causation is a form of substance-causation that does not involve an event, it provides no explanatory content and in fact involves causation without any mechanism. A person can cause an event, one might argue, but only as mediated by another event—for instance, a person can break a window, but only through the event of swinging the hammer. On some accounts of agent-causation, it is no better than an event completely driven by chance, since the absence of an event associated with the agent voids all explanatory content from agent-causation.[728] Van Inwagen objects that, since an agent-

[725] Henry Allison, *KTF*: 88. Kant makes the same point in *Gr* 4:454.
[726] Kant criticizes the idea that the mere existence of a substance, apart from mutual interaction, can account for change, in *ND* 1:413.
[727] Watkins notes Kant's use of the phrase "causality of the cause" to indicate this middle position between the substance and an event. Watkins (2005): 244, 249
[728] One can find this objection in Randolph Clarke (1993): 191.

cause lacks an event-cause, the agent is unable to control what it does.[729] Again, this objection rests on a monolithic notion of grounding. The presupposition of this objection is that the only possible ground of an action is an antecedently determining event—i.e. some mechanism whereby the agent exercises its causality. However, the agent-cause can contribute to an action, as its constitutive ground. Kant, in the *Prolegomena*, states that practical reason lies at the ground of appearances and is in fact the cause of the natural laws that characterize the empirical world.[730] The crucial distinction that is needed to understand how the agent-cause contributes to action without mediation by an event comes in Kant's resolution of the Fourth Antinomy, in which he distinguishes between a cause in appearances and a cause of appearances.[731] The agent-cause does not participate as a cause in appearances, since it is not an antecedent event. Those who reject transcendental idealism conclude that agent-causation is no better than chance, since they think the only way to constrain events through causation is through a cause in appearances. Those who embrace transcendental idealism understand that causation does not have to take place in appearances. There is a cause *of* appearances, in the sense of an absolute ground that is detached from human sensibility. The possibility of an intelligible ground of the time-series shows that causation without events can still have significant explanatory content.

Alfred Mele worries that an agent able to act differently, through an agent-causal power, even given an identical past and laws of nature, can only act arbitrarily.[732] But the

[729] This objection involves God rolling back time in repeated experiments and is discussed in Mele (2006): 23.
[730] *P* 346.
[731] *KdrV* A565/B593.
[732] Mele (2006): 69.

identity of the past, in my account, is complemented by a richly structured cause *of* appearances, i.e. an intelligible ground. The structure of the intelligible ground allows the agent not only to be the source of her action, but also provides explanatory resources independent of the identity of the past.

The notion of a cause of appearances can prevent the infinite regress some theorists insert into their accounts of agent-causation, in order to avoid this threat of chance. If an agent-cause produces event *q*, then, if *q* is not to arrive by chance, then there must be an event, *p*, in virtue of which the agent-cause produces *q*. At the same time, though, we want to make sure that event *p* does not also occur by chance. If we think only in terms of antecedent grounds, we have to posit another event between the agent and event *p*, and the infinite regress begins.[733]

Conclusion

In this chapter, I developed an account of the metaphysical framework for agent-causation, which involves a tiered metaphysics involving both an intelligible and empirical series, as well as both schematized and unschematized forms of contingency. I showed how this metaphysical framework has advantages over reductionist accounts of agent-causation, that articulate it as a psychological drive to act according to reasons, as

[733] There is a discussion of this issue in Timothy O'Connor (2000): 63. Chisholm actually endorses this infinite nesting of agent-causing in Chisholm (1971): 40-3.

well as accounts that rely on a dogmatic form of intuitionism. I then tested my account of agent-causation against both historical and contemporary objections.

One issue that Henry Allison introduces is that, by making the empirical series of an agent's behavior metaphysically dependent on the intelligible series of maxims, we take away any realism about this empirical character, making it into a sort of fraudulent image.[734] Karl Ameriks, however, debunks the idea that being metaphysically dependent entails complete subjectivism. Color, for instance, depends on the subjective constitution of the observer, and so does not characterize things in themselves. Nevertheless, insofar as it is grounded in the nature of things—things give off a certain color because of the way they reflect light—color has a level of objectivity that makes it more than a merely phenomenalistic entity.[735] Also, Allison is committed to a view of freedom in which we are free from the practical perspective, but which does not involve the idea that absolute freedom interacts with the empirical level as a real metaphysical ground. In spite of its metaphysical vacuity, we are supposed to act under the idea of freedom with rational necessity—it is more than a merely heuristic device we might adopt to better explain something. The idea of freedom is a necessary framework, for Allison, that governs the kind of explanation we look for in understanding our practical agency, which transcends merely passively caused empirical events.[736] Allison's lack of metaphysical qualification

[734] Henry Allison (2004): 46.
[735] Karl Ameriks makes this point in Karl Ameriks (2012a): 76. The ontological subjectivity of space and time does not entail their epistemological subjectivity. Space and time are epistemologically objective because they are necessary conditions of knowledge. Guyer (1979): 70. In the same passage, Guyer also makes reference to an attribution of objectivity to color in *KdU.* Color has objectivity in the sense that we expect others to see the same color we see.
[736] Allison distinguishes his own view of Kant from the intentional stance of Dennett, which is just a pragmatic device meant to help us understand and predict the behavior of certain kinds of being. Henry Allison (2012): 91.

of the empirical order, as ideal, means that reason's aspiration towards the unconditioned is potentially illusory, and the mere perspective of practical agency, rather than the possibility of the real metaphysical ground, is an unsystematized, causally inert, and indeed optional stance, since we can abandon this perspective for one from which we are determined.[737] Allison does not *metaphysically* qualify the empirical order, because space and time for him are understood only in terms of their epistemic function. The deliberative perspective therefore can only be entertained in an anti-realist manner that cannot be integrated into an equally valid third-person perspective in which we view ourselves as determined. For Kant, genuine scientific knowledge is an apodictic system, not based on mere empirical generalizations.[738] The presence of a perspective that is supposed to be necessary, alongside a whole body of theoretical cognition that is not properly adjudicated to accommodate this perspective, is not systematic knowledge.

However, if the intelligible character is the ground of what occurs in the empirical character, it seems as though conflict between these characters is impossible.[739] It seems a reasonable general principle that something cannot constrain the very ground of its existence, since upsetting its own ground would prevent it from existing. Yet, Kant continually refers to a conflict between reason's perception of the moral law and inclination. For instance, human nature is frail insofar as our incentive to follow the objective moral law is continually subverted by the subjective power of our selfish desires.[740] The idea, though, that a grounding relationship is incompatible with conflict

[737] Determinism, for Allison, is also true from a certain point of view with norms specific to it. Allison (2004):48.
[738] *MadN* 4:468.
[739] Paul Guyer made this objection in a talk given at the University of Edinburgh, on January 27, 2012, entitled "Hume, Kant, and the Passion for Reason."
[740] *R* 6:30.

presupposes a certain kind of conflict, i.e. one with a metaphysical structure that pits reason against inclination as two events. For a conflict to exist, one might argue, reason and inclination have arise as competing psychological events; reason qua ground complements the inclination, and cannot subvert it. However, the kind of conflict between reason and inclination I develop diversifies the metaphysical structure while preserving conflict. The intelligible character, as a ground (and so not as an event in time) develops a general orientation with respect to the incentives of either the moral law or non-moral incentives. This orientation is a supreme maxim that is not a specific event in time, but a temporally indeterminate orientation. This general orientation then privileges certain sets of reasons—either moral or non-moral ones—so that individual actions at specific times follow certain trajectories. So, there is conflict between reason and inclination, on my model, although not as an ontologically monolithic conflict between events, but in terms of a general orientation between two incentives that shapes individual actions at the empirical level.[741]

[741] See the ambiguity in the term "deed" in *R* 6:32. A deed can be a choice of a supreme maxim that governs one's will across time, or a specific choice in a situation arising at a particular instant in time.

CONCLUSION

AGENT CAUSATION WITHIN TRANSCENDENTAL IDEALISM

In this dissertation, I have developed an agent-causal account of freedom using the metaphysical framework of Kant's doctrine of transcendental idealism. The benefit of transcendental idealism comes mainly from the way it diversifies our understanding of grounding relationships. The diversification of grounding relationships associated with TI has to do with the introduction, in addition to the antecedent events in time that are part of the event-causal model, of a new ground in the form of a transcendental ground of the empirical series. In this conclusion, I will summarize and finalize my findings on the differences between, and implications of, a metaphysics of agency that involves only events, and one that, through the doctrine of TI, is able to involve a richer metaphysics of substances, powers, and grounds in which events are embedded. In the first section, I will describe a metaphysics of agency that appeals exclusively to events. In the second section, I will articulate the metaphysical principles underlying an agent-causal account that appeals to structures underlying and going beyond mere events. In the third section, I will show how Kantian metaphysics can unite the immanent and transeunt causation peculiar to Aristotle's opaque but foundational comments about agent-causation. In the fourth section, I will raise questions for future research.

Agency In the Event-Causal Series

Event-causal libertarianism presupposes a metaphysics that consists exclusively in antecedently determining events in time, and rejects the incorporation of temporally indeterminate grounds or powers. Event-causation is Humean in origin, insofar as it breaks down reality into temporally indivisible instants—i.e. events—that are represented in epistemologically primitive impressions.[742] Event-causation, in rooting its metaphysics of agency exclusively in what is epistemologically primitive, seeks a greater proportionality to empirical evidence than agent-causal theories. The idea of an agent-cause, acting as an underlying substance that directly produces its actions independently of event, is empirically obscure, insofar events represent the basic unit of empirical validation. Unlike Kant, the Humean ties ideas directly to empirical input. The simple impressions we receive constitute the fundamental boundaries of what is and is not a legitimate idea. Any cognitive state that goes beyond simple impressions is the beginning of what Hume calls "airy systems," i.e. metaphysical systems that have the air of sober science, claiming to give us insight into the ultimate springs of reality, but which are really deceptive and unjustified fabrications of mere speculation.[743] Agent-causation, with its invocation of substance-causation, offends against the Humean aversion to substances, which is based on the idea that there is some stable object underlying and unifying successive changes is a mere fiction.[744] The Kantian sees cognition as a union between intuitions and pure concepts of the understanding, and so rejects the Myth of the

[742] David Hume (1949): 170-218.
[743] David Hume (1999): 1.12.
[744] David Hume (1949): 220.

Given that the exclusive controlling factor for cognitive states are represented by empirical input.[745]

Agent-causation not only presupposes the idea of a substance, but also the idea of a power. The agent adjudicates between its different motivational states, in a decisive action that is not a merely indeterministic fallout. The sort of power an agent is supposed to exercise is inconsistent with reductive models of causality, which merely present events in patterns of regular association. An agential power cannot be subject to these patterns of regularity, since the agential power is able to produce its effects at will, in a deliberate selection between alternative possibilities.[746] Event-event models of causality, since they involve only fragmentary instants, cannot incorporate causal powers. Hume, of course, famously rejects the idea of causal powers, claiming that, when we observe one billiard ball strike another, we merely observe two events, and not some underlying power whereby the outcome of the collision is bound up with some faculty in the first billiard ball.[747]

Donald Davidson developed an influential event-causal account of action. He theorized that rational explanation of an action is a species of ordinary causal explanation. That is, we identify reasons for action by identifying its causes, in the form of some kind of pro-attitude (e.g. a desire) towards a certain action, as well as a belief that one is performing the action towards which one has a pro-attitude.[748] Action here is

[745] There is an account of the Myth of the Given in Eric Watkins (2008): 512-31.
[746] See G.E.M. Anscombe for a polemic against merely epistemological understanding of causation involving lawful relationships without some notion of real derivation of an effect from a cause through a power. G.E.M. Anscombe (1975): 67.
[747] David Hume (1949): Section VII.
[748] Davidson refers to this combination of a desire and a belief as a primary reason in Donald Davidson (2003): 4.

explained in terms of events in the form of the arising of mental states, and there is no appeal to underlying grounds—i.e. a ground that is not itself an event, but instead forms a backdrop necessary for the temporal determination of events. Davidson rejects agent-causation as a rationally unstable theory. Either the agent-cause produces an action by way of a mediating event, or there is an unmediated and direct relationship between the agent and its act, without a mediating event. The introduction of a mediating event means that the original "action" the agent supposedly caused is not really an action, but an effect of what the agent actually did, which was the mediating event that leads to the "action." The original "action" is no longer a primitive action—i.e. one with which the agent has a direct relationship—but is instead a remote effect of what the agent does. We then have a question about the mediating event: in virtue of what is it the product of agent-causation, except through another mediating event? We are then led to an infinite regress, in which we keep accounting for action in terms a mediating event of agent-causation of the agent. The alternative would be that there is no mediating event between an agent and its action, which means that the action simply arises, through mere happenstance. What is needed is the notion of an activity, possessed by the agent, and not the agent itself, which nevertheless is not a temporally determinate event.[749]

Events are not the only way to get explanatory content in the relationship between an agent and her act, and Davidson is actually showing a paucity of metaphysical imagination. For Davidson, actions have reasons in virtue of a causal relationship between events, and an attempt to establish a direct relationship between an agent-substance and an event is actually explanatorily opaque and non-rational. However, the

[749] Watkins is excellent on this issue. Watkins (2005): 260.

Kantian perspective, in which temporal determination itself requires grounds in what is temporally indeterminate, strips the privilege of rationality from explanation by events, since the temporal determination of events itself requires explanation. Rather than being limited to either mediating events that go on to entangle one in an infinite regress, or an irreducible and so explanatorily empty relationship between an agent and its action, the Kantian is able to appeal to the notion of an activity, whereby a substance is able to underwrite temporally specific events in virtue of its temporally indeterminate relationships with other substances.[750] The metaphysics of an ongoing activity characterizing a substance in mutual interaction with another substance—one might imagine, for instance, some formal principle involving a maxim that represents a continuous activity underlying different empirical manifestations of action—can provide the link to the agent by which an agent establishes control over her action without another event, leading to an infinite regress, or just a purely empty claim of direct substance-causation.

Metaphysical Structures Underlying the Kantian Agent-Causal Account

Kantian causality is not a mere relationship between determinations of things that occur at specific times—i.e. it does not adhere to the Humean rule of adhering strictly to

[750] These interrelationships between substances, in virtue of which substances can produce changing determinations in one another, go beyond the mere co-existence of substances. Kant uses the phrase, the "causality of the cause," to signify a special state in which substances are in mutual interaction, which surpasses their mere coexistence. In his pre-Critical period, Kant claims that substances stand in mutual interrelation with one another only in virtue of someone extrinsic to them, i.e. having a common cause in God. *ND* 1:413.

empirical input, and negating the idea of an irreducible power connecting events in favor of a reductionist psychological account of causation involving mental associations. We derive the outlines of a Kantian model of causality from his earliest works, specifically in two key principles of his *New Elucidation*, the principles of coexistence and of succession. The principle of succession holds that no change—i.e. no event—can occur unless there are substances in interaction with one another. A substance by itself is constrained by its internal grounds which prevent it from taking on opposing determinations. So, to stimulate movement beyond these internal grounds, the substance has to be placed in interaction with another substance. The principle of coexistence states that a substance by itself will not automatically enter into relationship with another substance—these relationships are extrinsic.[751] The principles of coexistence and succession together instantiate the Kantian phrase of the "causality of a cause."[752] The causality of the cause is distinct from the substance itself, since we know from the principle of coexistence that the substance by itself does not have an intrinsic activity by which it enters into relationships with other substances. The causality of the cause is also not an event, since it is an activity connecting substances in mutual interaction that is supposed to underwrite events. The causality of the cause is asymmetrical with events insofar as, according to the principle of succession, the activity of the causality of the cause binding substances together in mutual interaction is required to account for events, and so it cannot itself be an event lest an infinite regress begins.[753]

---

[751] Ibid., 1:413.

[752] See the sources for this phrase as cited in Eric Watkins (2005): 244.

[753] There is a passage in the argument for the Second Analogy in which Kant claims that the primary ground of appearances must not change, because this would require another ground. *KdrV* A205/B250. See Eric Watkins (2005): 263, note 40, for a precise explanation of why grounds are temporally indeterminate.

The Analogies of Experience in the First Critique continue the theme of the Principle of Succession: there can be no arising in time except against the backdrop of a substance of which the arising is a determination, and there can be new determination of a substance but for a cause in virtue of which the new determination takes its place in objective succession.[754] In turn, a substance cannot have its own place in time except against the backdrop of mutual interaction between substances. Causally isolated substances would not be able to ground an objective difference between simultaneity and succession, since the mere successive perception of two events cannot indicate whether the events are simultaneous or successive.[755]

All of this pre-Critical and Critical metaphysics may seem abstract, but it actually lends itself perfectly to agent-causal accounts of freedom. Just as agent-causal accounts of agency move beyond the exclusive reliance on events to articulate its metaphysical framework, the abstract elements in Kant's model of causality show how events rely on substructures that themselves are not localizable at a specific time. The causality of a cause is a ground of a substance exercising a continuous—and so not temporally localized—activity of mutual interaction with another substance. Causation for Kant does not consist in strings of events, each of which has temporal specificity. We can transpose this metaphysics into an agent-causal account smoothly, since the agent in agent-causation is supposed to represent something that is not temporally localizable. So, the agent is able to produce its own determinations, not as another event, but as a temporally indeterminate ground, just as a substance produces changes in the determinations of

[754] Kant discusses the Analogies of Experience at *KdrV* A182.
[755] *KdrV* A212/B259.

another substance in virtue of an activity uniting its grounds with those of another substance.

At the same time, the agent-substance does not produce determinations through its mere existence. If the mere existence of an agent-substance were to produce an event, we would have a problem of asymmetry, in which we would not be able to explain why something that is temporally non-specific—the mere existence is not rooted in a specific time—produces an event at a specific time. This asymmetry problem has plagued contemporary defenders of agent-causation, since it seems as though we need to appeal to an event to explain why action occurs at one time and not another.[756] Kant is clear, though, that it is not the mere existence of substances that accounts for time-determination. In order to mutually determine events in one another, substances have to have some sort of mutual interdependence that is based on relations that do not intrinsically follow from their mere existence. There is, then, an activity—not mere existence—of mutual interrelationships that makes for temporally specific events.[757]

This ongoing activity, that is not tied down to a specific moment in time, can take the form of an ongoing maxim that provides a structuring cause for several individual actions. The ongoing activity can act as a substantial form that imparts a structure to events in time: it does not function as itself an event leading to an action, but instead is a temporally non-specific form that imparts structure to a series of event-causal relationships. So, moving from the abstract to the concrete, one's maxim to be a good

[756] We find this objection, for instance, in defenders of event-causation like C.D. Broad and Donald Davidson. Cited in Eric Watkins (2005): 243.
[757] This is the Principle of Co-Existence, in *ND* 1:413. The idea that mutual interaction occurs by way of the mere existence of substances is a vulgar form of physical influx theory that Kant reject in *Diss.* Sect. 17.

student represents a kind of activity, uniting one in mutual interaction with the various interfaces of one's daily life, and manifesting itself in several different events: e.g. going to class, doing one's homework, going to office hours, etc.

We can see again the idea of an ongoing activity that is more metaphysically fundamental than individual events in Kant's *Metaphysical Foundations of Natural Science,* specifically in the discussion of attractive and repulsive forces between bodies. The exercise of attractive and repulsive forces continues even in a state of equilibrium. There are no changes of state in this balance of forces, but nevertheless there is ongoing activity.[758] The ongoing exercise of attractive and repulsive forces, in a state of equilibrium and so unaccompanied by any changes of state, evidences a grounding relationship that has the potential for explanatory content even without events. To return to the dilemma of Davidson about how the agent can have control over its action, we are not limited merely to another event or an explanatorily empty bald assertion. Instead, we can appeal to the notion of an activity that is not itself an event, which underwrites events and abides beyond them just as forces continue in interaction even when an object is not in motion.

The mathematical-mechanical mode of explanation, evident in event-causal accounts of action like that of Davidson, makes foundational the movement of bits of matter through the void, so that change is accounted for entirely in terms of the shape and movement of matter in space-time. For Davidson, events in space-time, with a token-token identity with physical things, are foundational in accounting for change. In the

[758] *MadN* 4:536-7.

metaphysical-dynamical mode of explanation, which Kant favors, fundamental moving forces of attraction and repulsion are foundational in accounting for change.[759] Again, the theme here is a metaphysical basis underwriting the empirical domain of events in space-time, which, for example, banishes the idea of absolute impenetrability, i.e. the idea that empirical objects fill space and resist other objects in virtue of their mere existence. On the contrary, Kant thinks that matter impenetrably fills space not in virtue of its existence, but in virtue of an underlying force.[760] If metaphysical forces underwrite specific events in the empirical domain, we can extend this metaphysics into our theory of free will, and incorporate an underlying activity, in virtue of which an agent can be the cause of its action without a mediating event.

Kant himself acknowledges that the mathematical-mechanical model of explanation—which appeals to a metaphysics along the lines of Democritus, which makes atoms the most fundamental ontological constituent—enjoys the advantage of being able to account for change in terms of the shapes of matter and empty space, which are both empirically quantifiable and observable. Yet, when we transpose the mathematical-mechanical model into a theory of agent-causation, we find a grave shortcoming in being unable to account for our deepest intuitions about ourselves. We are led, in fact, into a theoretical view of the world that is practically unsustainable, as we reduce persons to mere sequences of events from which their own agency is detached. The desires and beliefs which drive action for Davidson are not abstracted from a larger causal sequence which is alien to the agent, and so the Davidsonian picture of agency

[759] Ibid., 4:525.
[760] Ibid., 4:497.

leads us to see action as flowing indistinguishably from the larger series of natural causes. Such a view of agency, if taken to its conclusion and applied practically, will lead us to take the so-called objective attitude[761], in which we cease relating to persons in a participatory way and instead see agents as having their actions happen to them, in the same way that someone with a psychiatric disorder loses control and has their behavior happen to them.

Kantian Metaphysics Resolve the Riddle of Aristotelian Agent-Causation

Aristotle is the father of the agent-causal account. The distinction between immanent and transeunt causation is key to this account. Transeunt causation consists in a causal relationship between something that exercises causality and an effect that is independent of that which exercises causality. So, a rock breaking a window is a form of transeunt causation, insofar as the causality of the rock manifests itself in an effect that is external to its own substance. Immanent causation, on the other hand, occurs within the same substance exercising causality—it generates a new determination in the substance. Roderick Chisholm captures the distinction between immanent and transeunt causation in the distinction between doing and making happen. When an agent does something, she causally contributes to her own action by changing a determination of herself. The rock, when it breaks the window, does not do anything, since it does not actualize any new determinations in itself, but simply, through merely mechanical forces having to do with

[761] Peter Strawson (1974): 21.

laws of nature and its own physical transfiguration, makes the breaking of the window happen.[762]

Aristotle is famous for another passage, also one of the founding philosophical passages for agent-causation, in which he discusses a sequence of events that traces a root from effect to cause until it bottoms out in a person, which is the initiator of the sequence and so ends the cause-effect series. In Aristotle's sequences, a staff moves a stone, the staff in turn is moved by a hand, which in turn is moved by a person.[763] The person has a mysterious property, by which it is exempt from the conditioned-condition relationship leading back from the stone. Nothing in the causal series is unaccounted for except the person.[764] The motion of the stone does not occur of itself. It is made to happen, and the motion of the stone is not something it does, to use Chisholm's language. The event of the staff moving—this new determination of its state occurring at a specific time—is accounted for by the hand, and the motion of the hand is accounted for by the person. The question that arises, then, is how the person is able to be exempt from the need to account for its activity in terms of an antecedent ground? We could simply deny that the person is an initiator of activity, in order to avoid the problem of a seemingly unaccountable change of state of the person. But, of course, this leads us towards compatibilism, in which a person is subject to alien causation outside of herself, and there is little consolation in the fact that the person, through her desires and beliefs, is the most immediate cause of the action. If we arbitrarily cut off antecedent grounds around the

[762] Chisholm's view is discussed in Dorothy Emmett (1985): 76.
[763] Aristotle's founding quotes on agent-causation, as well as Chisholm's re-deployment of them, are cited in Dorothy Emmett (1985): 76.
[764] The action of the person is up to her in a special way, as an agent-cause independent of surrounding causal networks.

person, and thus break up the unity of experience, in which every event can be accounted for in terms of an antecedent ground, we still have serious problems in explaining how the action of moving the hand is rooted in the person. The person is not an event, and the problem of temporal asymmetry arises: how does a temporally indeterminate entity produce the temporally determinate motion of the hand? If we identify an event characterizing the person—some change in the state of the person by which she causes her hand to move—then we simply postpone the problem, as we then need to ask in virtue of what this mediating event is connected to the person. This is the path towards an infinite regress. If the agent is the mere subject of her action, and there is not some mediating event of causation, we face again the problem of non-causal libertarianism that I discussed in chapter 4: that is, without some kind of causal relationship to the action, the action arises of itself, and the agent has to be a spectator of the action, which has an intrinsic power onto itself.

Kant says, in his *Religion*, that "Man himself must make or have made himself into whatever, in a moral sense, whether good or evil, he is or is to become."[765] This self-making cannot occur if the agent is the product of deterministic causation external to her. If we try to establish indeterminism in order to create space for the agent to create herself, we end up with action that is product of chance—this is hardly the sort of self-making Kant has in mind. Self-making seems to require, moreover, an impossible infinite regress, since the agent has to continually reach back to control antecedent grounds. If an agent is to have rational control over an action, she has to have control over the grounds of the action, and if she is to have rational control over the grounds of action, she has to have to

[765] *R* 6:44.

control of the grounds of the grounds. Positing a contra-causal power, able to do otherwise in different possible worlds in spite of identical laws of nature and past events, seems to deprive this power of rational origination—i.e. differences in its exercise occur arbitrarily.[766]

We resolve these problems through the insight that it is not just antecedent grounds we are dealing with in accounting for the self-making of the person. The infinite regress of impossibility arguments against agency are resolved in light of the fact that agent can have rational control of her action, not through continually reaching back towards antecedent grounds, but through constituting herself via an intelligible ground. A contra-causal power of an agent-cause, able to act differently even given fact of an identical past, is not lacking in explanatory content, because constitutive-intelligible grounds provide an explanatory basis independently of the series of events in the past. Kantian metaphysics provides a special entry point by which we can deal with the otherwise mysterious stopping point we see in Aristotle's rendition of action, in which the person is held up as the only member of the causal sequence that is exempt from a rule tying it to a previous event. This special capacity for person to act as an unconditioned cause might be asserted in a non-philosophical flight, i.e. we could accept conditioned-condition relationships as universal throughout nature, but, in a leap of faith in which we surrender our reason, we could attribute a special sort of causality to the person.

[766] This is a concern of Mele in Mele (2006): 60.

Kant's approach is much more sophisticated than the non-philosophical flight. Transcendental idealism gives the person a depth that allows it to transcend the sequence of events in the empirical series. Transcendental idealism means that we do not have to mysteriously insert the person, as an unconditioned cause, as a member of the empirical series in time. This insertion would involve an arising in time for which there would be no accounting—it just happened, without any prior cause—and so it would involve a counterexample to the Second Analogy of Experience, which establishes a causal law as the necessary ground of objective succession in time. There is no need for such a violation, since transcendental idealism allows us to identify the empirical series as a conditioned reality, affiliated in particular with human sensibility. We can identify the unconditioned causality of the person, therefore, as a ground of the empirical series that is itself outside of time, and so not subject to the causal law.

Kant's TI unites immanent and transeunt causation. The transeunt causation leading up from the stone to the person moving her hand is event-causation in time, represented by events that are separate from each other transferring their power to a subsequent event. The activity of the person, on the other hand, is immanent, insofar as the person changes her own state independently of any antecedent event. We can exempt ourselves from event-causation in virtue of a transcendental cause of the empirical series, which itself is contingent and so in need of a metaphysically absolute ground. The immanent activity of the person represents a temporally indeterminate activity: it is not itself an event, but an ongoing activity structuring the person's interactions with her own body. Being the unconditioned cause of our action means that we can be both the source of our action, insofar as our unconditioned aspect acts out of the time-series and so

independently of the causal law, and also have leeway to select alternative possibilities, insofar as the lack of constraint by prior causes allows us to determine ourselves in different directions.

The unconditioned cause associated with the person is not in the empirical time-series. Instead, it plays a role similar to what Fred Dretske calls a "structuring cause." Dretske differentiates a process from a mere sequence of events. A sequence of events is what is available on a Humean view, in which the fundamental unit of ontology consists in disconnected impressions of slices of space-time. A process, in Dretske's technical usage of the term, is a sort of complex unity that is richer and transcends the mere events of which it is composed. In the same way a marriage is not reducible to the people who are in it, but instead is a special sort of unity that has its own independent properties: a marriage, for instance, can be annulled, whereas people cannot. Emphasizing that action is a process, which forms a causal unity between different events that has its own irreducible identity independent of the events, is a rejection of a view that reduces action to its output—i.e. the final product of what someone does. Action is not the intrinsic power of self-determination of mental properties we get in non-causal libertarianism, which is just output detached from any sort of process attaching it to the agent. We cannot isolate action in terms of events linearly leading to a product. Linear causal relationships between events characterize what are clearly non-actions—reflexive behaviors such a cockroach scurrying away when one approaches it. The structuring cause of one's character acts in a way analogous to the way the wiring of a doorbell system functions: the reason why pressing the doorbell itself causes a ringing sound (this is the linear causal relationship between an event and an output) is because of the internal

wiring of the doorbell. Structuring causes are not part of the linear relationship between two temporally distinct events, but part of a temporally indeterminate (because the wiring remains while different events of button-pressing come and go) backdrop to this linear relationship.[767] In a similar manner, the metaphysics of agency I have developed in this dissertation, while acknowledging linear relationships between antecedent events and their effects, also appeals to freely chosen maxims of an individual's character. These maxims are temporally indeterminate, like wiring for the doorbell. They also account for why certain psychological events lead to certain manifestations of action, and so function as structuring causes. Structuring causes ensure that action is a process: in does not consist in just constituent events, but a causal unity that is irreducible to individual events. Action is a process that is irreducible to individual, temporally determinate events: it involves an interplay between these events and temporally indeterminate structural factors that account for the connection between events.

Ideas for Future Research

A possible avenue for future research could involve exploring the connection between a metaphysical power of absolute freedom and Kant's aesthetics, particular his idea of sublimity. The sublime is contrasted with the beautiful insofar as the latter is characterized by purposiveness, so that it has a natural and satisfying fit with our power of judgment, whereas the former is characterized by a contrapurposiveness that goes

[767] See Fred Dretske (1988): Chapter 2.

against the categorical structure immanent to experience. The sublime cannot be contained in any sensible form. Absolute freedom, of course, cannot be an object of possible experience, since it violates the causal unity essential to experience. Sensible forms can only symbolize the ideas of reason associated with the sublime, which transcend any possible sensible embodiment. For instance, the raging sea, which itself conforms to the immanent a priori structure of experience, provides nevertheless a symbol of uncontained violence that leads the mind to consider the ideas of reason that transcend all experience.[768] Also of interest in the third *Critique* is the idea that there is a chasm sealing off the domain of the supersensible containing absolute freedom, and the theoretical cognition of nature, but for the mediating faculty of judgment. Judgment connects the lawfulness of nature with the final end represented by the innate dignity in the idea of freedom, in the form of the concept of the purposiveness of nature. Nature is purposive, that is, because its laws are directed towards the final end represented by freedom.[769] Freedom seems to play a role, then, in both sublimity, since it transcends sensible forms, as well as beauty, since freedom represents the final end around which the purposiveness of nature is structured. The third *Critique* therefore provides a gateway by which we might understand the relationship between freedom and aesthetic phenomena, as well with the teleological structure of the world.

A second possible avenue of research could be in the philosophy of mind. I have articulated an immaterialist metaphysics in this dissertation, and so I am opposed to an

[768] *KdU* 5:246. Henry Allison notes that a key thesis of the first *Critique* is that reason aims for a totality that the understanding can never capture. Allison (2001): 319.
[769] *KdU* 5:195.

identity theory that collapses the mental into mere physical states of the brain.[770] It would be interesting to see to what extent my theory of free will can be integrated into a theory such as the pan-psychism of David Chalmers, which claims that consciousness is an intrinsic property of physical entities. Consciousness, that is, is a categorical basis for the relational properties we see in the physical world, and which we study through the discipline of physics. So, a quark consists in a set of relational properties by which it interacts with other substances, but physics cannot give an account, in addition to these relational properties, of what the quark is in itself. Chalmers's view is that, intrinsically, reality is phenomenal, i.e. conscious.[771] Obviously, my view of agency is similar, insofar as events in the empirical series are grounded in a non-spatiotemporal level involving merely thinkable (not experienceable as things embodied in space-time) and formally structuring maxims. The immaterialist metaphysics I introduce is similar to Chalmers's pan-psychism in its disavowal of the fundamentality of the spatiotemporal physical realm.

A third area of future research I will mention has to do with the unity or heterogeneity of the notion of grounding.[772] I mention in this dissertation a diversification in grounding relationships from Kant's pre-Critical to his Critical periods. It will be interesting to pursue, in light of contemporary controversies over the potential unity of different kinds of grounding relationships, whether there is a unified way to understand grounding both in time, and as a transcendental cause.

[770] For a defender of mind-brain identity, see J.J.C. Smart (2002): 60.
[771] See David Chalmers (2002): 265.
[772] Jonathan Schaffer mentions the controversy over the univocity of the grounding relationship in Jonathan Schaffer (2016): 51.

The fourth and final area of future research stemming from this dissertation I will mention has to do with a tension between Kant's apparent commitment to strict deterministic lawfulness at the level of human psychology in the first *Critique*, and his denial in the *Metaphysical Foundations* that empirical psychology can be a science.[773] Kant says in the latter work that empirical psychology, which tries to establish laws for the phenomena of "inner sense," has a lesser epistemic status than even chemistry, which Kant designates as a "systematic art," rather than proper science, because it cannot incorporate a priori intuitions that would represent a mathematical basis for chemistry.[774] Psychology is still more inchoate, because humans are much less amenable to experimentation than chemical components and observation of inner sense destabilizes the phenomenon under observation so as to skew results.[775]

Kant's attitude to empirical psychology is to be distinguished from transcendental psychology, in which we study the faculties necessary for cognition. Empirical psychology studies passing inner phenomena in time—the changing inner theater of desires, fears, idle reflections, etc.[776]—and not the a priori nature of the various faculties, such as the understanding, reason, sensibility, that lie at the root of cognition. Empirical psychology is also to be distinguished from rational psychology, insofar as rational psychology hypostasizes the formal structures necessary for cognition in transcendental psychology and treats them as a noumenal spiritual substance beyond possible experience—i.e. it is not an empirical discipline.[777]

---

[773] The denial occurs at *MadN* 4:471.
[774] Ibid.,
[775] 4:471.
[776] *KdrV* A347.
[777] *KdrV* A348/B406.

It is strange that, given Kant's rejection of the idea that empirical psychology can be a science, in the first *Critique*, Kant states that "if we could investigate all the appearances of his [i.e. a human being] power of choice down to their basis, then there would be no human action that we could not predict with certainty…"[778] It needs to be noted that Kant has very high standards for identifying a discipline as scientific, since, for him, genuine natural science is ensconced in an a priori framework of both metaphysics and mathematics. Metaphysics identifies the pure structure of the categories of the understanding which impart a general structure of nature independently of any object. Mathematics takes empirical concepts and articulates them in terms of a priori intuitions, e.g. the empirical phenomena of motion can be expressed as geometrical or numerical relationships. It could be the case that, though empirical psychology cannot be expressed mathematically, it is nevertheless subject to strict lawfulness permitting, in theory, perfect prediction.

[778] *KdrV* A549/B577.

BIBLIOGRAPHY

Adams, Robert M. *Leibniz: Determinist, Theist, Idealist*. New York: Oxford University Press, 1994.

Adams, Robert M. "Things in Themselves." *Philosophy and Phenomenological Research*, Vol. 5, no. 4, (1997): 801-825.

Allais, Lucy. "Kant's One World: Interpreting 'Transcendental Idealism." *British Journal for the History of Philosophy*, Vol. 12, no. 4 (2004): 655-684.

Allais, Lucy. *Manifest Reality*. Oxford: Oxford University Press, 2015.

Allison, Henry. *The Kant-Eberhard Controversy*. Baltimore: The Johns Hopkins University Press, 1973.

Allison, Henry. "Kant on Freedom: A Reply to My Critics." *Inquiry*, 36 (1993): 443-64.

Allison, Henry. "Kant's Refutation of Realism." *Dialectica*, 30 (1976): 223-253.

Allison, Henry. *Kant's Theory of Taste*. Cambridge: Cambridge University Press, 2001.

Allison, Henry. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 2000.

Allison, Henry. *Kant's Transcendental Idealism.* New Haven: Yale University Press, 2004.

Allison, Henry. "Kant's Refutation of Materialism." *The Monist*, Vol. 72, no. 2 (1989): 190-208.

Allison, Henry. “On Naturalizing Kant’s Transcendental Psychology.” *Dialectica*. Vol. 49, no. 4 (1995): 335-356.

Allison, Henry. “On the Presumed Gap in the Derivation of the Categorical Imperative.” *Philosophical Topics*. Vol. 19, no. 1 (1991): 1-15.

Allison, Henry. “The Non-Spatiality of Things in Themselves in Kant.” *Journal of the History of Philosophy*, 14, no. 3 (1976): 313-321.

Allison, Henry. “Reflections on the Banality of (Radical) Evil: A Kantian Analysis.” *Graduate Faculty Philosophy Journal*, 18, no. 2 (1995): 141-158.

Allison, Henry. “Transcendental Realism, Empirical Realism, and Transcendental Idealism.” *Kantian Review*, Vol. 11 (2006): 1-28.

Allison, Henry. “We Can Act Only Under the Idea of Freedom.” In *Essays on Kant*. Oxford: Oxford University Press, 2012: 87-98.

Ameriks, Karl. “From Kant to Frank: The Ineliminable Subject.” In *The Modern Subject: Conceptions of the Self in Classical German Philosophy*. Edited by Karl Ameriks and Dieter Sturma. Albany: State University of New York Press, 1995: 217-227.

Ameriks, Karl. *Interpreting Kant’s Critiques*. Oxford: Oxford University Press, 2003.

Ameriks, Karl. *Kant and the Fate of Autonomy*. Cambridge: Cambridge University Press, 2000.

Ameriks, Karl. *Kant and the Historical Turn.* Oxford: Clarendon Press, 2006.

Ameriks, Karl. *Kant’s Elliptical Path*. Oxford: Clarendon Press, 2012a.

Ameriks, Karl. “The Question of Whether a Purely Apparent Person is Possible.” In *Spinoza and German Idealism*. Edited by Eckart Forster and Yitzhak Y. Melamed. Cambridge: Cambridge University Press, 2012: 44-58.

Ameriks, Karl. “Review: Idealism and Freedom: Essays on Kant’s Theoretical and Practical Philosophy, by Henry Allison.” *Philosophy and Phenomenological Research*. Vol. 59, no. 3 (1999): 825-829.

Anscombe, Elisabeth. “Causality and Determination.” In *Causation and Conditionals*. Edited by Ernest Sosa. Oxford: Oxford University Press, 1975: 63-81.

Anscombe, Elisabeth. *Intention*. Cambridge, MA: Harvard University Press, 1957.

Armstrong, David. *A World of States of Affairs*. Cambridge: Cambridge University Press, 1999.

Audi, Paul. “Grounding: Toward A Theory of the In-Virtue-Of Relation.” *Journal of Philosophy*, Vol. 109, no. 12 (2012): 685-711.

Balaguer, Mark. *Free Will As An Open Scientific Problem*. Cambridge, MA: MIT Press, 2010.

Baumgarten, Alexander. *Metaphysics*. Translated and edited by Courtney D. Fugate and John Hymers. London: Bloomsbury, 2013.

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

Bird, Graham. *Kant’s Theory of Knowledge*. London: 1962.

Boehm, Omri. “Kant’s Idea of the Unconditioned and Spinoza’s: The Fourth Antinomy and the Ideal of Reason.” In *Spinoza and German Idealism*. Edited by Eckart

Förster and Yitzhak Y. Melamed. Cambridge: Cambridge University Press, 2012: 27-43.

Brewer, Kimberly and Eric Watkins, "Difficulty Still Awaits: Kant, Spinoza, and the Threat of Theological Determinism." *Kant-Studien*, Vol. 103, no. 2 (2012): 163-187.

Carnap, Rudolf. "Replies and Systematic Expositions." *The Philosophy of Rudolf Carnap*. Edited by Paul Schlipp. La Salle: Open Court, 1963: 859-867.

Chalmers, David. "Consciousness and Its Place in Nature." In *Philosophy of Mind: Classical and Contemporary Readings*. Edited by David Chalmers. Oxford: Oxford University Press, 2002: 247-272.

Chisholm, Roderick. "Freedom and Action." In *Freedom and Determinism*. Edited by K. Lehrer. New York: Random House, 1966: 11-44.

Chisholm, Roderick. *Person and Object*. La Salle, IL: Open Court, 1976.

Chisholm, Roderick. "Reflections on Human Agency." *Idealistic Studies*, Vol. 1, (1971): 33-46.

Clarke, Randolph. *Libertarian Accounts of Free Will*. New York: Oxford University Press, 2003.

Clarke, Randolph. "Towards a Credible Agent-Causal Account of Free Will." *Nous* Vol. 27, no. 2 (1993): 191-203.

Davidson, Donald. "Actions, Reasons, and Causes." *Journal of Philosophy*. Vol. 60, no. 23 (1963): 685-700.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980.

Dretske, Fred. *Explaining Behavior*. Bradford: MIT Press, 1988.

Dummett, Michael. *Truth and Other Enigmas*. London: Duckworth, 1978.

Emmett, Dorothy. *The Effectiveness of Causes*. Albany: State University of New York Press, 1985.

Enc, B. *How We Act: Causes, Reasons, and Intentions*. Oxford: Oxford University Press, 2003.

Ertl, Wolfgang. "Ludewig Molina and Kant's Libertarian Compatibilism." In *Companion to Luis de Molina*. Edited by Matthias Kaufmann and Alexander Aichele. Leiden: Brill, 2014: 405-446.

Euler, Leonhard. "Letters to a German Princess." In *Kant's Critique of Pure Reason: Background Source Materials*. Edited and translated by Eric Watkins. Cambridge: Cambridge University Press, 2009: 183-230.

Fine, Kit. "The Notion of Ground." In *Metaphysical Grounding: Understanding the Structure of Reality*. Edited by Fabrice Correia and Benjamin Schneider. Cambridge: Cambridge University Press, 2012: 37-80.

Forster, Michael N. "Herder and Spinoza." In *Spinoza and German Idealism*. Edited by Eckart Forster and Yitzhak Y. Melamed. Cambridge: Cambridge University Press, 2012: 59-84.

Franklin, Christopher. "Farewell to the Luck (and Mind) Argument." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. Vol. 156. No. 2 (2011): 199-230.

Garber, Daniel. “Leibniz and the Foundations of Physics: The Middle Years.” In *The Natural Philosophy of Leibniz*. Edited by Kathleen Okruhlik and James Robert Brown. Dordrecht: D. Reidel, 1985: 27-130.

Gardner, Sebastian. “Review: Kant’s Transcendental Idealism: An Interpretation and Defense, Revised and Enlarged Edition.” *Notre Dame Philosophical Reviews*, 2005.09.04.

Goetz, Stewart. *Freedom, Teleology, and Evil*. New York: Continuum, 2008.

Ginet, Carl. *On Action*. Cambridge: Cambridge University Press, 1990.

Goldman, A. *A Theory of Human Action.* Englewood-Cliffs, NJ: Prentice Hall, 1970.

Goyette, John. “St. Thomas on the Unity of Substantial Form.” *Nova et Vetera*, Vol. 7, no. 4 (2009): 781-790.

Griffiths, Meghan. “Freedom and Trying: Understanding Agent-Causal Exertions.” *Acta Analytica*, Vol. 22, no. 1 (2007): 16-28.

Grier, Michelle. *Kant’s Doctrine of Transcendental Illusion*. Cambridge: Cambridge University Press, 2001.

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

Paul Guyer. *Kant and the Claims of Taste*. Cambridge, MA: Harvard University Press, 1979.

Guyer, Paul and Ralph Walker. “Kant’s Conception of Empirical Law.” *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 64 (1990): 221-258.

Hatfield, Gary. “Kant on the Perception of Space (and Time).” In *Kant and Modern Philosophy*. Edited by Paul Guyer. Cambridge: Cambridge University Press, 2006: 61-93.

Henrich, Dieter. *Between Kant and Hegel*. Cambridge, MA: Harvard University Press, 2003.

Henrich, Dieter. “The Proof-Structure of Kant’s Transcendental Deduction.” *The Review of Metaphysics*, Vol. 22, no, 4 (1969): 640-659.

Hogan, Desmond. “Noumenal Affection.” *Philosophical Review*, Vol. 118, no. 4 (2009): 501-532.

Hudson, Hud. *Kant’s Compatibilism*. Ithaca: Cornell University Press, 1994.

Hume, David. *An Enquiry Concerning Human Understanding*. Edited by Tom L. Beauchamp. (Oxford: Oxford University Press, 1999).

Hume, David. *The Treatise of Human Nature*. La Salle, Illinois: The Open Court Publishing Company, 1949.

Jacobi, F.H. “David Hume über den Glauben: oder Idealismus und Realismus.” In *Werke*. Leipzig: Gerhard Fleischer, 1815.

Janiak, Andrew. *Newton As a Philosopher*. Cambridge: Cambridge University Press, 2008.

Jauernig, Anja. “Kant’s Critique of the Leibnizian Philosophy: Contra the Leibnizians, But Pro Leibniz.” In *Kant and the Early Moderns*. Edited by Daniel Garber and Beatrice Longuenesse. Princeton: Princeton University Press, 2008: 41-63.

Kane, Robert. *The Significance of Free Will*. Oxford: Oxford University Press, 1996.

Kant, Immanuel. *Critique of the Power of Judgment*. Edited and translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press, 2000.

Kant, Immanuel. "Critique of Practical Reason." In *Practical Philosophy*. Edited and translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kant, Immanuel. *Critique of Pure Reason*. Edited and translated by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, 1998.

Kant, Immanuel. "Groundwork of the Metaphysics of Morals." In *Practical Philosophy*. Edited and translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.

Kant, Immanuel. "Inaugural Dissertation." In *Theoretical Philosophy 1755-1770*. Edited and translated by David Walford and Ralf Meerbote. Cambridge: Cambridge University Press, 2002.

Kant, Immanuel. "Fourth Letter: From Kant to Herz, February 21, 1772." In *Kant's Critique of Pure Reason: Background Source Materials*. Edited and translated by Eric Watkins. Cambridge: Cambridge University Press, 2009: 312-315.

Kant, Immanuel. *Metaphysical Foundations of the Natural Sciences*. Translated and edited by Michael Friedman. Cambridge: Cambridge University Press, 2004.

Kant, Immanuel. "A New Elucidation of the First Principles of Metaphysical Cognition." In *Theoretical Philosophy 1755-1770*. Edited and translated by David Walford and Ralf Meerbote. Cambridge: Cambridge University Press, 2002.

Kant, Immanuel. “On A Discovery According to Which Any New Critique of Pure Reason Has Been Made Superfluous By An Earlier One.” In *The Kant-Eberhard Controversy*. Edited and translated by Henry Allison. Baltimore: The Johns Hopkins University Press, 1973.

Kant, Immanuel. *Prolegomena to Any Future Metaphysics*. Translated and edited by Peter G. Lucas. Manchester: Manchester University Press, 1953.

Kant, Immanuel. *Religion Within the Boundaries of Mere Reason*. Translated and edited by by George di Giovanni and Allen Wood. Cambridge: Cambridge University Press, 1998.

Kant, Immanuel. “True Estimation of Living Forces.” In *Natural Sciences*. Edited and translated by Eric Watkins. Cambridge: Cambridge University Press, 2012.

Kant, Immanuel. “What Does It Mean to Orient Oneself in Thinking?” In Immanuel Kant, *Religion Within the Boundaries of Mere Reason*. Edited and translated by Allen Wood and George di Giovanni. Cambridge: Cambridge University Press, 1998: 1-14.

Kim, Jaegwon. “The Many Problems of Mental Causation.” In *Philosophy of Mind: Classical and Contemporary Readings*. Edited by David Chalmers. Oxford: Oxford University Press, 2002: 170-178.

Kim, Jaegwon. “Multiple Realizability and the Metaphysics of Reduction.” *Supervenience and Mind*. New York: Cambridge University Press, 1993: 309-35.

Korsgaard, Christine. “Morality As Freedom.” In *Kant’s Practical Philosophy Reconsidered.* Edited by Yirmiyahu Yovel. Boston: Kluwer, 1989: pgs. 23-48.

Korsgaard, Christine. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.

Langton, Rae. *Kantian Humility*. Oxford: Clarendon Press, 1998.

Leibniz, G.W. *The Leibniz-Clarke Correspondence*. Edited and translated by H.G. Alexander. New York: Manchester University Press, 1956.

Leibniz, G.W. *Monadology and Other Philosophical Essays*. Translated and edited by Paul Schrecker and Anne Martin Schrekecker. Pearson: Library of the Liberal Arts.

Lewis, David. “Humean Supervenience Debugged.” *Mind* 103 (1994): 473-90.

Lewis, David. *Philosophical Papers 2*. New York: Oxford University Press, 1986.

Locke, John. *An Essay Concerning Human Understanding*. Edited by Kenneth P. Winkler. Indianapolis: Hackett Publishing Company, 1996.

Longuenesse, Beatrice. “Kant’s Multiple Concepts of Person.” *Kant on Persons and Agency*. Edited by Eric Watkins. Cambridge: Cambridge University Press, 2017. 155-173.

Loux, Michael. *Primary Ousia: An Essay on Aristotle’s Metaphysics Z and H*. Ithaca, NY: Cornell University Press, 2018.

Manley, David. “Introduction: A Guided Tour of Metametaphysics.” *Metametaphysics: New Essays on the Foundations of Ontology*. Edited by David Manley, David Chalmers, and Ryan Wassermann. Oxford: Oxford University Press, 2009: 1-37.

McTaggart, J. Ellis. “The Unreality of Time.” *Mind*, 17, no. 68 (1908): 457-474.

Mele, Alfred. *Free Will and Luck*. New York: Oxford University Press, 2006.

Melnick, Arthur. *Kant’s Analogies of Experience*. Chicago: University of Chicago Press, 1973.

Melnick, Arthur. “Review of *Kant’s Transcendental Idealism*.” *The Philosophical Review*, Vol. 94, no. 1 (1985): 134-137.

Miller, Kristie. “Presentism, Eternalism, and the Growing Block.” In *A Companion to the Philosophy of Time*. Malden, Mass.: Wiley-Blackwell, 2013: 345-364.

Mumford, Stephen. *Dispositions*. Oxford: Oxford University Press, 1998.

Nelkin, Dana. *Making Sense of Freedom and Responsibility*. Oxford: Oxford University Press, 2011.

Nelkin, Dana. “Two Standpoints and the Belief in Freedom.” *The Journal of Philosophy*, Vol. 97, no. 10 (2000): 569.

Newton, Isaac. “De Gravitatione.” In *Isaac Newton: Philosophical Writings*. Edited by Andrew Janiak. New York: Cambridge University Press, 2004.

O’Connor, Timothy. “Emergent Properties.” *American Philosophical Quarterly*. Vol. 31, no. 2 (1994): 91-104.

O’Connor, Timothy. *Persons & Causes*. Oxford: Oxford University Press, 2000.

O'Neil, Onora. "Four Models of Practical Reasoning." In *Bounds of Justice*. Cambridge: Cambridge University Press, 2000: 11-28.

O'Neil, Onora. "Kant After Virtue." *Inquiry* 26 (1983): 387-405.

Ostaric, Lara. "Kant's Account of Nature's Systematicity and the Unity of Theoretical and Practical Reason." *Inquiry*, Vol. 52, no. 2 (2009): 153-178.

Ostaric, Lara. "Practical Cognition, Reflective Judgment, and the Realism of Kant's Moral *Glaube*." In *Realism and Anti-Realism in Kant's Moral Philosophy: New Essays*. Edited by Robinson dos Santos and Elke Elisabeth Schmidt. Band: De Gruyter, 1999: 91-118.

Derk Pereboom. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press, 2014.

Pereboom, Derk and Michael McKenna. *Free Will: A Contemporary Introduction*. New York: Taylor & Francis, 2016.

Pereboom, Derk. *Living Without Free Will*. Cambridge: Cambridge University Press, 2001.

Posy, Carl. "Dancing to the Antinomy: A Proposal For Transcendental Idealism." *American Philosophical Quarterly*, Vol. 20, no. 1 (1983): 81-94.

Price, Huw. "Carnap, Quine, and the Fate of Metaphysics." *Electronic Journal of Analytic Philosophy*, Vol. 5, no. 1 (1997): 1-22.

Putnam, Hillary. *Reason, Truth, and History*. Cambridge: Cambridge University Press, 1998.

Rosenberg, Gregg. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. Oxford: Oxford University Press, 2004.

Royce, Josiah. *The Spirit of Modern Philosophy*. Boston: Houghton Mifflin Company, 1892.

Schaffer, Jonathan. "Grounding in the Image of Causation." *Philosophical Studies* 173 (2016): 49-100.

Rumore, Paola. "Kant's Understanding of the Enlightenment With Reference to His Refutation of Materialism." *Contextos-Kantianos: International Journal of Philosophy*. No. 1 (2014): 81-98.

Schaffer, Jonathan. "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*. Edited by David Manley, David Chalmers, and Ryan Wassermann. Oxford: Oxford University Press, 2009: 347-383.

Stang, Nicholas. "The Non-Identity of Appearances and Things in Themselves." *Nous* 48, no. 1 (2014): 106-136.

Sider, Theodor. "Presentism and Ontological Commitment." *The Journal of Philosophy*, Vol. 96, no. 7 (1999): 325-347.

Smart, J.J.C. "Sensations and Brain Processes." In *Philosophy of Mind: Classical and Contemporary Readings*. Edited by David Chalmers. Oxford: Oxford University Press, 2002: 60-67.

Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies*, 75 (1994): 5-24.

Spinoza, Benedictus. *Ethics*, trans. by Edwin Curley. Princeton: Princeton University Press, 1985.

Strawson, Peter. *The Bounds of Sense*. New York: Routledge, 1966.

Strawson, Peter. "Freedom and Resentment." *Proceedings of the British Academy*, 48 (1962): 187-211.

Strawson, Peter. "The Problem of Realism and the A Priori." In *Kant and Contemporary Epistemology*. Edited by Paolo Parrini. Dordrecht: Springer Netherlands, 1994: 167-173.

Timpe, Kevin. *Free Will 2nd Edition: Sourcehood and Its Alternatives*. London: Bloomsbury Academic, 2013.

Van Cleve, James. *Problems from Kant*. New York: Oxford University Press, 1999.

Velleman, J. David. "What Happens When Someone Acts?" *Mind*, Vol. 101, no. 403 (1992): 461-481.

Watkins, Eric. "Kant on Extension and Force: Critical Appropriations of Leibniz and Kant." In *Between Leibniz, Newton, and Kant: Philosophy and Science in the Eighteenth Century*. Dordrecht: Springer, 2001: 111-127.

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

Watkins, Eric. "Kant and the Myth of the Given." *Inquiry*, Vol. 51 (2008): 512-531.

Watson, Gary. “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme.” In *Perspectives on Moral Responsibility*. Edited by John Martin Fischer and Mark Ravizza. Ithaca: Cornell University Press, 1993: 119-148.

Wolff, Robert Paul. *Kant’s Theory of Mental Activity*. Cambridge, MA: Cambridge University Press, 1963.

Westphal, Kenneth. “Noumenal Causality Reconsidered: Affection, Agency, and Meaning in Kant.” *Canadian Journal of Philosophy*, Vol. 27, no. 2 (1997): 209-245.

Wittgenstein, Ludwig. *On Certainty*. Edited by G.E.M. Anscombe and G.H. von Wright. Oxford: Blackwell, 1969.

Wolf, Susan. *Freedom Within Reason*. Oxford: Oxford University Press, 1990.

Wood, Allen. “Kant’s Compatibilism.” *Self and Nature in Kant’s Philosophy*. Edited by Allen Wood. Ithaca and London: Cornell University Press, 1984: 73-101.