

15 Metaphysics on the Model of Natural Science?

A Kantian Critique of Abductivism

Nicholas Stang

15.1 Introduction

For several decades in the 20th century, analytic philosophy suffered under the despotic rule of logical empiricism, and metaphysics was driven into exile. But, as Kant predicted, the reign of the empiricists proved to be brief, and metaphysics was returned, if not to the throne, at least to a central place in contemporary philosophy. Work on metaphysical topics that would have been familiar to Kant's rationalist predecessors – e.g., monism, grounding, the principle of sufficient reason – now thrives in analytic philosophy, while the once-feared principle of verification, long used to drive metaphysics into hiding, finds few adherents.

In an ironic reversal, though, those who defend metaphysics today often do so by claiming that it is continuous with natural science. The very accusation that Kant thought undermined metaphysics – that 'the birth of the purported queen [of the sciences] was traced to the rabble of common experience' (Ax) – is now taken to vindicate it instead.

This is the view of metaphysics I will call *abductivism*. Abductivism is a package of views, a key component of which is a thesis about the common methodology of metaphysics and science:

Abductive Methodology. The basic methodology of metaphysics is the same as the basic methodology of the sciences, i.e., inference to the best explanation (IBE, or abduction for short).¹

Just as the physicist infers the theory that best explains, for instance, the paths of particles in a cloud chamber, so too, according to the abductivist, does the metaphysician infer the theory that best explains some target phenomenon, such as the modal profiles of objects, or their persistence through time. Abductivism is also committed to a thesis about the epistemology of metaphysics:

Abductive Epistemology. The core source of our knowledge of metaphysical theories is abduction.²

Of course, the abductivist need not think that *all* of our knowledge in metaphysics comes from abduction; some of our knowledge in metaphysics might be knowledge of logic, knowledge of conceptual truths, or knowledge of mathematics or physics. But that knowledge by itself is not enough to get us knowledge of metaphysical theories, and the difference is made up by abduction. The empirical, conceptual, logical, mathematical, and physical truths under-determine metaphysical theory, so we infer to the metaphysical theory that best explains the “data”.

Another component of abductivism is realism about natural science:³

Scientific realism. The theoretical terms in our best scientific theories refer to entities and structures in the world that are (i) theory-independent and (ii) mind-independent.

Since abductivists see metaphysics as continuous with science, they are committed to the parallel claim about metaphysics:

Metaphysical realism. The theoretical terms in our best metaphysical theories refer to entities and structures in the world that are (i) theory-independent and (ii) mind-independent.

Abductivism is thus committed, in Kantian terms, to *transcendental realism*.⁴

Abductivism is an attractively unified set of views about the methodology, epistemology, scientific status, and semantics of metaphysics. It has many contemporary adherents.⁵ Abductivism offers a putative explanation of the possibility of metaphysics as knowledge and as a science. In Kantian terms, abductivism is a *critique of metaphysical reason*, an explanation of how (abductive) reason can achieve knowledge in metaphysics that simultaneously establishes the limits of reason in metaphysics.^{6,7} Reason can go as far in metaphysics as abduction reaches; once our inferences to the best explanation gives out, our rational warrant for metaphysical theorising gives out too, and we should refrain from speculating further.⁸

In fact, Kant agrees with abductivism on a few core claims. He agrees that metaphysics and natural science are about explanations, i.e., not merely knowing what is the case, but also knowing *why*.⁹ He further agrees with the abductivist about the first conjuncts of *Scientific realism* and *Metaphysical realism*, though he rejects the second conjunct: the concepts in our natural scientific and metaphysical theories refer to *theory*-independent entities and structures in the world which are *mind-dependent*.¹⁰

But this narrow band of agreement conceals a vast gulf of disagreement about the very nature of metaphysics itself. For Kant, abduction provides only *comparative* certainty in its conclusion: the grounds of the inference (the data) make the conclusion (the putative explanation) probable (or at least more probable than the alternatives), but they are not sufficient

guarantors of its truth; one can abduct from incomplete or misleading data to a false conclusion.¹¹ But metaphysics, according to Kant, must have *apodictic* certainty: the grounds on which we base our metaphysical claims must suffice, by themselves, to ensure that those claims are true.^{12,13}

Because he has this apodictic conception of metaphysics, Kant must reject both *Abductive methodology* and *Abductive epistemology*. Kant thus cannot accept an abductivist conception of metaphysics. But neither can he accept a purely abductivist conception of natural science. Kant thinks that all natural sciences, properly speaking, rest on metaphysical, hence apodictically certain, foundations, which therefore cannot be justified abductively.¹⁴ Once these metaphysical foundations are in place, though, inference to the best explanation can play a significant role in natural science.¹⁵

Given this vast difference in their respective conceptions of metaphysics, prospects for an informative dialogue between Kant and the abductivist seem bleak. In this chapter, I try to break this impasse by exploring Kant's reasons for thinking that abductivist metaphysics cannot do what it claims to do, that is, it cannot explain the possibility of knowledge in metaphysics. But because the abductivist conception of metaphysics is so radically unlike Kant's conception of metaphysics, and in fact is more similar to his conception of (the empirical part of) natural science, it will actually be *Kant's critique of natural science* that will be relevant to abductive metaphysics. I will argue that Kant's arguments about the possibility of natural science can be redeployed to show that abductive metaphysics fails. What is more, these arguments do not depend on Kant's controversial assumption, which would in any case be rejected by any contemporary metaphysician, that metaphysics must be apodictically certain.

Although abduction plays a significant role in the metaphysical theories of some of his predecessors¹⁶ and in his own pre-Critical metaphysics,¹⁷ Kant was not acquainted with any philosopher who held to the strictly abductivist conception of metaphysics.¹⁸ Consequently, my manner of proceeding in this chapter will be more reconstructive than strictly exegetical. Although this chapter is based on a reading of Kant developed elsewhere, it is not *primarily* a work of Kant exegesis but an attempt to take some ideas from Kant and use them to criticise contemporary metaphysics. Nonetheless, I will indicate along the way how my argument maps onto Kant's texts.

15.2 Abductive Metaphysics

How is abductive knowledge in metaphysics, knowledge by inference to the best explanation, possible? To answer this question, we need to say a bit more about what *explanation* is. There is a vast literature on this topic, and it is not my intention to address it comprehensively here.¹⁹ Instead, I

will attribute a fairly simple (possibly oversimplified) set of commitments about explanation to the abductivist. Over the course of this chapter, I will consider whether the abductivist can evade my arguments by abandoning or revising some of these commitments.

Let us begin by distinguishing two senses of *explanation*. Firstly, we can understand explanation as an activity of epistemic agents, and derivatively as the product of that explanation (i.e., the activity of explaining produces explanations). Secondly, we can understand explanations as entities or structures in the world, the real things that “back” or “correspond to” our explanations in the first sense. For instance, causal explanation in the first sense is the activity of searching for, and finding, causes for given events; causal explanations in the second sense are the causes themselves, the causes referred to and described by explanation in the first sense. It is controversial whether the activity of explanation needs to be “backed by” or “correspond to” worldly explanatory structure, but this assumption is shared by many contemporary metaphysicians, so I will take it for granted.²⁰

The first condition, therefore, that must be satisfied for the possibility of metaphysical knowledge, as the abductivist conceives it, is that the world have explanatory structure. Furthermore, it must have the kind of explanatory structure relevant to metaphysical explanations. If metaphysical explanation includes causal explanation, then there must be causal structure in the world (i.e., there must be causes and effects). Likewise, if metaphysical explanation involves various kinds of non-causal explanations, then it must include relevant corresponding worldly explanatory relations. For instance, the possibility of *grounding explanation* in metaphysics (i.e., explaining facts by citing their *grounds*) requires that there be grounding structure in the world. Which particular explanatory structures a given metaphysician is committed to will depend on the details of her metaphysical theory; the key idea is that abductivist metaphysicians are committed to there being ontic explanatory structure “in the world”.

According to the abductivist, the core source of our knowledge of metaphysical theories is abduction.

Since knowledge requires truth, the abductivist is committed to holding that abduction in metaphysical theories is *truth-tracking*, i.e., it generally takes us from true premises to true conclusions. But for abduction to be successful, it not only needs to be *alethically* successful (truth-tracking), it must also be *explanatorily* successful, i.e., it must take us from true premises to truths *about their explanations*. These are distinct requirements; a mode of inference (e.g., logical deduction) might be truth-tracking without necessarily tracking explanations of its premises. Finally, if abduction in metaphysics is to provide *knowledge*, then its alethic and explanatory success cannot be a matter of *mere luck*. “Mere luck” here is simply a stand-in for the post-Gettier idea that knowledge is incompatible with being simply

a matter of accident, of having “gotten lucky” in having true beliefs. In other words, abduction in metaphysics must provide (i) *un-Gettier-able* (i.e., non-lucky) (ii) *true* beliefs about the (iii) *explanations* of the premises.

But the distinction drawn above between *explanation-as-something-we-do* and *explanation-in-the-world* commits the abductivist even further. For our practices of explanation have a certain structure, sometimes called “canons of theory choice” or “canons of explanation”: explanations must be unified, parsimonious, rich in consequences, agree with other theories, etc. These features are what constitute the *bestness* of inference to the *best* explanation. For example, to pick a feature of explanations that almost everyone agrees is a virtue, the abductivist is committed to claiming that *unified* explanations are *ceterus paribus* better than disunified ones. But in order for abduction guided by this principle to be “alethically successful” there must be unified ontic explanatory structure “in the world”. If there were, for instance, fundamentally distinct kinds of explanatory structure in different domains (or even different regions of the universe), then in abducting to the more unified explanatory theory we would not be tracking the ontic explanatory structure, i.e., we would not be making explanatorily successful inferences (even if they were alethically successful). Note that the *ceterus paribus* character of our canons of theory choice means that the requirements imposed on the ontic explanatory structure in the world are somewhat “loose”: the ontic explanatory structure must be to some degree unified, fruitful (rich in consequences), etc.²¹

Successful explanation (and thus successful IBE) requires more than there just being explanatory structure in the world. Explanations are sensitive to structure. If I explain some target phenomena by citing an explanans that gets the structure of the underlying ontic explanation wrong, I have failed to explain the phenomena, even if my beliefs about the explanans are true. This means that the success of IBE is sensitive to the *language* in which we conduct it.²² If our language does not “carve” ontic explanatory structure “at the joints”, then our explanations of facts stated in that language, assuming the explanations are also formulated in our language, will refer to non-joint-carving entities and structures in the world. We may systematically infer to true theories, but they will not be theories backed by the objective explanatory structure in the world. Our inferences may track the *truth*, but they will fail to be *successful* inferences to the best explanation.

I will illustrate these claims with a pair of thought experiments. I will then draw a general lesson from them.

Recall the famous “grue” example from Goodman 1944 (updated for 2023): there is a community whose language includes the predicate “grue”, which applies to objects if and only if they are *either* green and first observed before 2050 *or* blue and first observed after 2050. Speakers of this language will infer in 2023 from the fact that all observed emeralds

are grue to the conclusion that all emeralds whatsoever are grue. This example is unsuitable for my purposes, for three reasons: it is a case of induction, not abduction; it is clearly not a case of metaphysical abduction; and the conclusion is false, rather than merely non-explanatory. To make my point I need a Goodman-style case involving metaphysical abduction to a true but non-explanatory conclusion.

15.2.1 *Gerrymandered Persistence*

Let us assume that perdurance is the correct explanation of why objects persist through time (i.e., exist at more than one time): objects persist by having instantaneous temporal parts at the times at which they exist. Objects, let us assume, are 4D spacetime “worms”: mereological fusions of instantaneous temporal parts. Consider a community that has correct beliefs about the patterns of persistence of objects, i.e., correct beliefs about which objects exist at which times, but that has a term in their language “gr-sists”, which refers to objects that persist *either* by perduring (having instantaneous temporal parts) *or* by enduring (by being “wholly present” at each time at which they exist, or however you want to spell out endurantism). They have no term in their language that refers specifically to the property of perduring or to the property of enduring. The speakers of this community will reason abductively as follows:

- (P1) [A set of true beliefs about the patterns of persistence of objects.]
 (C1) ∴ The gr-sistence theory of the persistence of objects.

By hypothesis, their abductive premise(s) and conclusion are true; the speakers of this community have true beliefs about the pattern of persistence of objects, and objects do gr-sist. But the conclusion does not explain why objects persist the way they do. They persist because they *perdure*. In fact, objects gr-sist because they perdure; perdurance is one of the disjuncts of gr-sistence, and disjuncts ground disjunctions (or so I will assume). The speakers of this community will infer to abductive conclusions that fail to explain their premises because their language includes the predicate “gr-sists”, rather than “perdures”. Their abductive inferences will fail because the language they speak fails to carve the world at its explanatory joints.

15.2.2 *Reversed Grounding*

Let us assume that there is an ontic explanatory relation in the world that answers to the contemporary notion of “grounding” and that standard assumptions about that relation are true, i.e., that it is irreflexive (nothing

grounds itself), asymmetric (no circular grounding), and transitive. Let us assume further that the actual world is *grounding paradise*: the Principle of Sufficient Reason holds for the grounding relation, so absolutely every fact has a ground (I will leave it open whether the PSR itself has a ground). Assume additionally that every fact has a consequence: every fact grounds some further fact (e.g., a disjunction of which it is one of the disjuncts). Consider a community whose language contains the term “gr-grounds” that refers to the inverse of the grounding relation: “*p* gr-grounds *q*” is true in their language just in case the fact that *q* grounds the fact that *p*.²³

Speakers of this language will infer from facts to their consequences, not to their grounds. Because every fact has a consequence (we have assumed), they may systematically infer to true conclusions, but they will systematically infer to the wrong (metaphysical) explanations; they will infer consequences, not grounds, as explanations. Consider, further, how such a community would construct a metaphysical theory of “what gr-grounds what” (to borrow Schaffer’s (2009) phrase). The premises of their abductive inference would consist in the facts that constitute the target phenomena of such a theory, while their conclusion would consist in a set of claims that gr-ground those facts. For instance, their theory of what gr-grounds the existence of material parts will cite facts about the existence of wholes of which they are parts (assuming that material objects are grounded in their parts); their theory of what gr-grounds the existence of complex material objects will cite facts about their shadows (assuming that the existence of complex material objects grounds the existence of their shadows), etc. In general, their metaphysical theorising will be alethically successful but not explanatorily successful, that is, it will not track ontic explanatory structure.

These examples are meant to constitute an intuitive case for the principle that theories will be explanatorily successful, i.e., track not only the truth but also the ontic explanatory structure, only if they involve terms that refer to ontic explanatory structure in the world rather than, for example, gerrymandered relations and properties.

To summarise, then, the Abductivist is committed to *all* of the following:

- A1. *Ontic explanatory structure*. There is enough ontic explanatory structure in the world to “back” explanations in true metaphysical theories.
- A2. *Abductive epistemology*. Abduction in metaphysics according to our canons of explanation (unity, etc.) tracks the truth and the underlying ontic explanatory structure, and both in a non-Gettierable (not merely lucky) fashion.
- A3. *Reference dependence*. The terms in our best metaphysical theories carve the ontic explanatory structure at its joints.

If A1 is false, then abductive metaphysics is impossible because there is not enough explanatory structure in the world to “back” our explanations. If A2 is false, then abduction in metaphysics might generate true beliefs, but it would not be knowledge (it would be accidentally true belief). And if A3 is false, abduction in metaphysics might generate knowledge, but it won’t generate knowledge of underlying explanatory structure; we will systematically fail to carve the world at its explanatory joints.

15.3 Kant’s Critique of Natural Scientific Abduction

Kant gives a *critique* of abduction in natural science, in his specific semi-technical sense of “critique”: an account of the nature and limits of a capacity for knowledge, in this case, our capacity for knowledge of Nature through inference to the best explanation. He does not offer a separate critique of abduction in metaphysics because he thinks metaphysics cannot be abductive; metaphysics, he thinks, must have apodictic, rather than merely comparative, certainty.²⁴ Furthermore, Kant’s critique of abduction, the bulk of which occurs in the Appendix to the Transcendental Dialectic, depends on his own transcendental idealist metaphysics, which by that point in the *KrV*, he takes himself to have proven.

Therefore, in using that Kantian critique to criticise contemporary abductive metaphysics I will have to proceed somewhat indirectly. In this section, I summarise Kant’s critique of abduction, organised around his explanations of principles in his own theory that correspond to A1–A3. Because my primary aim here is systematic rather than textual, I will simply state what I take Kant’s view to be, keeping textual details, and my reasons for reading Kant the way I do (as well as engagement with secondary literature), to a minimum.²⁵ The aim of this section is twofold: to motivate a critical question about abductivism (what explains A1–A3?) and to give some reasons for thinking of this as a genuinely Kantian question, the kind of question Kant (or “the Kantian”) should ask when confronted with abductive metaphysics.

I will formulate these three Kantian theses K1–K3 in the contemporary terms of A1–A3, in order to bring out their similarities. This will involve some translation of Kant into contemporary lingo; Kant would not formulate them this way exactly (nor would I if my task were primarily exegetical).

K1. *Ontic explanatory structure.* There is enough ontic explanatory structure in the spatiotemporal world to back explanations in natural scientific theories.

The key to Kant’s explanation of K1 is his transcendental idealism, the doctrine that the form of the spatiotemporal world is grounded in the form

of our experience of the spatiotemporal world. Transcendental idealism entails the following:

Transcendental idealism. If our experience, in virtue of its very form, represents the spatiotemporal world as having enough ontic explanatory structure to back natural scientific theories, then this grounds the presence of sufficient such structure in the spatiotemporal world.²⁶

Because the spatiotemporal world is a world of appearances, and the “being” of appearances is grounded in the content of our experience of them, if we experience the world as having ontic explanatory structure (e.g., causal structure), this grounds the presence of that structure in the world of appearances. But notice that *Transcendental idealism* will explain the presence of structure only in the phenomenal world, the world of appearances, not in things in themselves. It can provide no explanation of why there is any ontic explanatory structure in things in themselves (assuming there is any). Kant’s explanation of the possibility of abduction is an explanation of the possibility of abduction from (premises about) appearances to (conclusions about) appearances. It provides no explanation of how we could use abduction to acquire knowledge of things in themselves, because Kant thinks such an abduction could never provide us with any knowledge. One immediate consequence is that Kant’s transcendental idealist explanation of K1 cannot be borrowed by the contemporary abductivist, who, in Kantian terms, is a transcendental realist about metaphysics.

Kant thinks that the principal kind of explanation in natural science is *causal* explanation, so the first thing he must account for is the presence of causal structure in the phenomenal world. He does so by arguing, in the Analogies of Experience, that:

Experience. Experience, in virtue of its very form, represents spatiotemporal objects as absolutely persisting substances in law-governed causal interaction, i.e., every alteration in a substance is the effect of a substance (whose state is reciprocally altered by the prior substance), where these alterations are governed by universal and necessary laws.

This, combined with *Transcendental idealism*, explains why the spatiotemporal world has causal and nomic structure: experience represents it as having this structure, and that grounds the presence of such structure (because the objects in questions are appearances).

But Kant thinks that the phenomenal world has further “explanation-backing” structure, which is crucial for his critical reconstruction of abduction in the Appendix to the Transcendental Dialectic (henceforth, “the Appendix”): real essences. The connection with causal-nomic structure is

this: laws are necessary in virtue of being grounded in the real essences of the objects about which they are laws.²⁷ This means that causal-nomic structure brings with it essence structure. Put more precisely, the connection is as follows:

Law. If it is a law that ϕ has ψ then (a) it is true that all samples of ϕ have ψ , and (b) every sample of ϕ has ψ in virtue of the *real essence* of ϕ .^{28,29}

Real essences are essences of (kinds) of things, unlike logical essences, which are essences of concepts.³⁰ Logical essences contain the marks that constitute a concept, while real essences contain the properties that constitute objects themselves.³¹ To use one of Kant's favourite examples, both attractive and repulsive forces are part of the real essence of matter, but only the latter is a mark contained in the concept *<matter>*.³² This means that it is a law that matter has both attractive and repulsive forces. Different speakers can associate the same empirical concept with different logical essences (they can give it different *nominal* definitions) but they refer to the same kind of object, e.g., matter.³³ Logical essences of concepts are relatively trivial for natural scientific purposes; they contain only the marks that speakers happen to think in a certain concept. What is far more important are the properties contained in the real essence, for they ground the laws that are the proper topic of scientific inquiry.^{34,35}

To use a piece of non-Kantian terminology, I will say that, when ϕ is as above, the ϕ s constitute a *natural kind*, or, equivalently, that *< ϕ >* is a *natural kind concept*. Some care is required here, because Kant is a conceptualist about universals: all that exist are individuals (individual substances and their individual accidents), and any "generic" or "universal" entity (e.g., a general property, a kind) "exists" only in the content of conceptual representation.³⁶ So natural kinds are not part of Kant's inventory of what exists, not part of what Quine would call his "ontology". Consequently, talk about natural kinds always has to be, in principle, paraphrasable in different terms: saying of a concept that it is a natural kind concept, or that objects fall under that concept in virtue of their sharing a common real essence.³⁷

Empirical concepts are at least partly individuated by the natural kinds to which they refer: if C and C^* are the same empirical concept, they refer to the same natural kind.³⁸ But natural kinds are themselves individuated by their real essences (even if we can never know the complete real essence).³⁹ The properties in the real essence make that kind the thing it is; without them, it would not be possible. For instance, nothing could be matter that did not have its essential attractive and repulsive forces. But if natural kinds themselves partly individuate the empirical concepts that refer to them, and natural kinds are themselves individuated by their

real essences, then real essences partly individuate empirical concepts. For instance, nothing could be the empirical concept *<matter>* if it did not refer to matter, i.e., a kind with the same real essence as matter (including attractive force).⁴⁰ This means that if there were no natural kinds whatsoever (constituted by objects sharing a generically similar real essence), then there would be no empirical concepts whatsoever. What unifies the instances of an empirical concept is not possession of the marks contained in its logical essence, for possession of these marks is not by itself sufficient to fall under the empirical concept; it is possession of the marks that constitute the real essence of the corresponding kind.⁴¹ For instance, possession of repulsive force is not sufficient to fall under *<matter>*; this requires attractive force as well, even though only the former, but not the latter, is contained in the logical essence of that concept. If there are no common real essences, there would be nothing to ground membership in a common empirical concept. In particular, mere perceptible similarity among objects is not by itself sufficient for them to fall under a common empirical concept; a generically similar real essence is required.

One immediate consequence of Kant's essentialist theory of empirical concepts and laws is that there is an isomorphism between the system, if there is one, of empirical concepts (a hierarchy of more and less general empirical concepts) and the system, if there is one, of natural laws (a hierarchy of more and less general laws of nature). Empirical concept C is more general than empirical concept C* if and only if the real essence of the natural kind corresponding to C is more general than the real essence of the natural kind corresponding to C*.⁴² But there is also an isomorphism between empirical natural kinds and natural laws: natural kind *k* is more general than natural kind *k** if and only if the *k* law is more general than the *k** law. By transitivity, this entails that there is an isomorphism between the system of empirical concepts and the system of natural laws: empirical concept C is more general than empirical concept C* if and only if the corresponding C law (the *k* law, where C is the concept of *k*) is more general than the corresponding C* law (the *k** law, where C* is the concept of *k**). We can translate unproblematically between a system (hierarchy of relative generality) of empirical concepts and a system of natural laws.

K2. *Abductive epistemology.* Abduction in natural science according to our canons of explanation (i.e., generality, specificity, continuity) tracks the truth and the underlying ontic explanatory structure, and both in a non-Gettierable (not merely lucky) fashion.

Despite the centrality of causal explanation to Kant's conception of natural science, and his recognition that many of our causal inferences will be

abductive (for a single effect could have many causes, which means we must infer to the most likely cause),⁴³ his most sustained discussion of abduction in natural science, in the Appendix, is not about causal inference, but inference to the system of empirical concepts (ordered into species and genuses). This is because Kant thinks a great deal more can be determined *a priori* about the causal laws that govern material substance than is given by the pure transcendental principles of the *KrV* alone (this project is undertaken in *MAN*). Only once these impure yet *a priori* “metaphysical” principles are in place is it possible to engage in causal inference about objects of experience.⁴⁴ Consequently, my focus in discussing K2 will be on natural scientific abduction that reveals the underlying system of empirical concepts, rather than on casual abduction.

Kant attributes the cognitive task of inferring from perceptible similarities and differences among objects to the underlying system of empirical species and genuses to what he calls the “hypothetical” use of reason. The hypothetical use of reason is clearly not deductive, ‘that is, not such that if one judges in all strictness the truth of the universal rule assumed as a hypothesis thereby follows’ (A647/B675).⁴⁵ But it is equally clear that the hypothetical use of reason is not exclusively (or even primarily) inductive. One of the main topics of the Appendix is the hypothetical use of reason in inferring from particular empirical concepts (and their associated laws) to more general empirical concepts, i.e., from species to genus. The inference from two or more species to their common genus is not an inductive inference, i.e., an inference from a premise of the form that *p* has held in all observed instances (e.g., that all observed samples are F) to the conclusion that *p* holds in all cases whatsoever (that all samples whatsoever are F). It is an inference from two or more empirical concepts to the common underlying features that unite *and (partially) explain* them, as well as potentially other species of the same genus. In contemporary terms, it is abductive.

Kant’s discussion of the principles that govern the hypothetical use of reason is complex, confusing, and controversial. What follows is a simplified presentation, sufficient for my purposes in this chapter.

Kant argues for three regulative principles that guide our abductive search for the system of empirical concepts: generality (For any two species, seek a common genus), specification (For any genus, seek further species), and continuity (Between any two species of a genus, search for an intermediate species). Together these constitute the regulative Idea of *systematicity* in Nature. In each case, Kant argues that these regulative principles (which he frequently dubs “logical”) are possible only under the assumption of a corresponding transcendental one. I will focus on generality: ‘The logical principle of genera therefore presupposes a transcendental one if it is to be applied to Nature (by which I here understand only objects

that are given to us). According to that principle, sameness of kind is necessarily presupposed in the manifold of a possible experience (even though we cannot determine its degree *a priori*), because without it no empirical concepts and hence no experience would be possible' (A654/B682). Corresponding claims are made about the logical/regulative principles of specification and continuity: they are applicable only on the assumption of a corresponding transcendental principle, without which experience would not be possible.⁴⁶

On my reading, the transcendental status of the principle of generality is a consequence of Kant's essentialist conception of empirical concepts: if there were no common real essences (if no empirical object were similar to any other object in virtue of their real essences), then there would be no empirical concepts.⁴⁷ The presence of common, perceptually manifest, qualities in objects is not sufficient to ground their belonging to the same empirical concept. The most this will guarantee is a common logical essence, formed by abstracting from their perceptible similarities; but, as we have seen, having the marks contained in a logical essence is not enough to ground membership in a common empirical concept. For that, common real essences are required. It is a commitment of this reading that the species and genera of which Kant speaks in the Appendix are essential species and genera: they divide objects, not according to their accidental or merely perceptually manifest properties but according to their real essences. Without empirical concepts, experience is impossible;⁴⁸ so without at least some common real essences, experience is impossible.⁴⁹ This is why the principle of generality is transcendental, i.e., it makes experience possible.

This is a sketch of Kant's explanation of K2 restricted to a very general formal feature of our explanations and the ontic structure that backs it: abduction in natural science according to the principle of generality (search for common empirical genera) must be backed by ontic explanatory structure constituted by natural kinds sharing common real essences. I say that this is general and formal because the form of experience explains why the search for common genera in general will not be entirely in vain. But it does not explain why our actual attempts to find these common genera will succeed. In particular, it leaves open the possibility that although Nature satisfies the principle of generality, indeed may even satisfy it ideally (for any two objects there is a common genus that subsumes both), those common genera are cognitively inaccessible to human beings. To make this vivid, it might be that the empirical similarities in terms of which we initially classify objects, due as they are to the contingent constitution of our sense organs (e.g., we classify green things as similar because of the contingent structure of our eyes), are systematically at variance with the underlying system of empirical genera, so that proceeding

from the former and trying to infer the latter, as we do, will never succeed. More generally, Nature might be ideally systematic (i.e., ideally satisfy the principles not only of generality, but also of specificity and continuity), but either so complicated or so unlike the system of perceptible similarities by which we initially categorise objects that we will never be able to achieve knowledge of the underlying system of species and genres.

This is the problem that Kant introduces in the two introductions to the *KU* and for which he introduces the regulative principle of purposiveness: Inquire into Nature *as though* it is purposive for our cognition, i.e., exists in order for us to cognise it.⁵⁰ Kant takes this to be equivalent to: Inquire into Nature as though it is the sensible product of non-sensible (non-spatiotemporal) intelligent author.⁵¹ In particular, if we assume this regulative principle, we will assume that our inferences to the species-genus system underlying Nature will, at least under the right conditions, track the truth (we will infer to true conclusions) and ontic explanatory structure (because the genres are real explanatory structures in the world), in a non-Gettierable fashion (we will not merely be getting lucky, for Nature was created, in part, to enable us to make these inferences in a reliable way).

This is Kant's explanation, such as it is, of K2. It raises difficult questions, principal among which is this: Can a regulative principle, like the principle of purposiveness, explain why K2 is true, or merely why we are rationally warranted in assuming K2 in our inquiry into Nature? But given my purposes in this chapter, I will leave that issue.⁵²

K3. Reference dependence. The concepts in our best natural scientific theories carve the ontic explanatory structure at its joints.

First, we need to distinguish, among the concepts in our natural scientific theories, between *a priori* formal concepts such as *<cause-effect>*, *<substance>*, *<law>*, and *<real essence>*, and empirical concepts of particular natural kinds of substances, their real essences, and the causal laws they obey (e.g., *<matter>*, *<water>*, *<gold>*, etc.). Kant's explanation works very differently in the two cases.

In the formal case, the ontic explanatory structure of the phenomenal world is grounded in the structure of our experience of that world, according to *Transcendental idealism*. Since these formal concepts (*<cause-effect>*, *<substance>*, *<law>*, and, I have argued, *<real essence>* as well) are part of that *a priori* form, they ground that structure in the phenomenal world. For instance, since the principle of generality is transcendental, there can be no experience that does not represent there being at least some real essence shared by some spatiotemporal objects; consequently, there is at least some such real essence shared by some spatiotemporal objects. But, by the same token, formal concepts represent the very structure

they ground, or in contemporary terms, they “carve” that structure “at its joints”. Given that the formal structure of our experience grounds the formal structure of the phenomenal world, there is no possibility that concepts in which that former structure is articulated will fail to correspond to the joints in the latter structure.

But this formal joint carving does not, by itself, entail that empirical specifications of these formal concepts will refer, much less that they will carve at the explanatory joints. It might be, for instance, that while our concept of *<natural kind>* carves at the joints (Nature is divided into natural kinds), none of our empirical concepts of specific natural kinds (e.g., *<water>*) refer, or, insofar as they do, they refer to gerrymandered kinds that do not carve at the joints (like *<grue>*). This gap is meant to be filled by the principle of purposiveness: if Nature exists for the sake of being comprehended by us, then, at least under the right conditions, our empirical natural kind concepts (and concepts of corresponding laws) will refer and carve at the joints. It may be that, at any given time, some of our empirical natural kind concepts fail to refer or carve at the joints, but, in the fullness of time, these will be replaced by empirical concepts that do carve at the joints. We are rationally warranted in inquiring into Nature as though this is the case, so we are rationally warranted in inquiring into Nature as though K3 is true. As with K2, whether this constitutes an explanation of why K3 is true, or merely of our rational warrant for assuming it in inquiry, remains unclear. But, again, rather than address that question directly, I will instead move on to consider how to mount a Kantian critique of abductive metaphysics.

15.4 A Kantian Critique of Abductive Metaphysics

The previous section established that Kant has at least the beginnings of an explanation of K1–K3. But this means that there is a natural Kantian critique of abductivism, beyond the mere insistence (rejected in any case by the abductivist) that metaphysics must have apodictic certainty, which of course abduction can never give us. Kant (or “the Kantian”) can ask, What is the abductivist explanation of A1–A3?

I think we can immediately dispense with one answer, which at least some abductivists will want to give: whatever explains the possibility of abduction in general explains A1–A3.⁵³ The reason that this answer by itself does not suffice is that, according to abductivism itself, metaphysics is more general than natural science.⁵⁴ Since metaphysics is more general, it is at least logically possible that abduction is alethically and explanatorily successful in natural science, but not in metaphysics.⁵⁵ It might, for instance, be the case that abduction is successful when we restrict our abductive premises and conclusions to natural scientific phenomena and

theories, but that it systematically misfires when we extend our inferences to the topics of metaphysics. For instance, it might be that the criteria by which we judge the “bestness” of an explanation in natural science does not apply in metaphysics, so that when we abduct by those criteria in metaphysics, our inferences systematically go wrong. In a Kantian vein, the fact that metaphysics since the time of Aristotle has been ‘a battlefield of endless controversies’ (Aviii) might even be taken as positive evidence that abduction does indeed fail in metaphysics, that when we extend our inquiry beyond the bounds of natural science we are left without any sufficient criteria for theory choice, and all we have is a ‘mock combat’ (Bxv).

With respect to A1, to my knowledge, no contemporary abductivist metaphysician attempts to explain *why* the world includes ontic explanatory structure necessary to “back” their explanatory theories. This is no accident. For, assuming that explanations must themselves be backed by explanatory structure, any explanation we might give of one ontic explanatory structure would require another one, *ad indefinitum*.⁵⁶ The dominant view among abductive metaphysicians is that at this point we reach explanatory “bedrock” and we cannot go any further.

Matters are slightly more complicated with A2. Again, I’m not aware of any abductive metaphysician who attempts to explain *why* A2 is true. The most that abductive metaphysicians offer is a “good company” argument: we adopt similar canons of explanatory reasoning (e.g., Occam’s razor) in physics and the rest of natural science as we do in metaphysics, so whatever explains why these are truth- and explanation-tracking in a non-lucky (non-Gettierable) fashion in the former case presumably does so in the latter as well.⁵⁷ But, as I argued in [Section 15.2](#), A2 is false without A3: unless our language carves the world at its explanatory joints, our abductive inferences will not generate explanatory knowledge. Furthermore, A1–A3 as claims about abduction in metaphysics are logically separable from corresponding claims about abduction in natural science. Assuming that abduction in natural science is alethically and explanatorily successful, neither entails that, nor explains why, abduction in metaphysics is alethically or explanatorily successful. It might be, for all that contemporary abductivists have shown, that we come to explanatory knowledge of natural phenomena through abduction in natural science, but these abductive methods systematically fail to produce knowledge when applied to metaphysics.

In the case of A3, there is a prominent philosopher, Ted Sider, who does purport to explain why the terms in our metaphysical theories carve the world at its joints. Sider (2011) gives a theory of reference in general, but, unlike many other reference theorists, he also explicitly applies his account to metaphysical reference.⁵⁸ Sider argues that the world has an objectively

privileged structure and the aim of metaphysics is to uncover that structure. He further holds that explanations must be couched in terms that carve that structure at its joints; of two putative explanations, the one that carves closer to the objective structure of the world is the better explanation. “Structure” in his account is thus playing a role similar to the role played by “ontic explanatory structure” in mine: the very same structure that “backs” explanations is also the structure we aim to uncover in metaphysics.

If we idealise somewhat and call the conjunction of all the sentences we hold to be platitudes or quasi-definitional of our terms our “theory”, the first part of Sider’s view is straightforward: any model of our theory (any assignment of objects to singular terms and extensions to predicates on which the theory is true) is a possible interpretation of the theory.⁵⁹ However, as Putnam (1981) pointed out, for any such model with a domain larger than a single object, there is a “permuted” model, which is also an interpretation of the original theory.⁶⁰ In general, these permuted interpretations will be grossly gerrymandered; they will assign intuitively bizarre objects and extensions to the singular terms and predicates of our language. In virtue of what is the intended interpretation of our theory the correct one? How do we eliminate, in a principled fashion, these gerrymandered models?

Sider follows Lewis (1984) in thinking that what determines reference is that some reference candidates are intrinsically more eligible for reference. Whereas Lewis originally restricted the notion to objects and predicates (some objects are more natural than others, some predicates pick out more natural properties than others), Sider extends the notion (which he calls being ‘structural’) to any item of any syntactic category whatsoever: quantifier domains, operators, relations of any adicity, even the logical connectives themselves can be said to be more or less structural. Among, the possible interpretations of our theory, the correct interpretation maximises structuralness.⁶¹ Reference is determined by descriptive fit plus structuralness.

Sider’s “official” notion of structure is an absolute one: an item is either structural or it is not.⁶² But reference cannot be determined by descriptive fit plus absolute structure, for that would entail the absurd result that we only ever refer to the absolutely structural, i.e., the fundamental metaphysical structure of reality. Not only would this make most of ordinary thought and talk impossible, it would also make Sider’s meta-semantics impossible, for it would make it impossible for us to refer to our own language, which, for Sider, is not itself absolutely structural. (The English language, plausibly enough, is not part of the fundamental structure of reality.) Thus, Sider’s view is that reference is determined by descriptive fit plus *comparative* structuralness. However, Sider never gives us more

than an intuitive sketch of what makes one item more structural than another. At one point he suggests an account of comparative structuralness in terms of the length of metaphysical definitions: *s* is more structural than *s** if and only if the metaphysical definition of *s* in absolutely structural terms is shorter than the definition of *s**.⁶³ But Sider immediately rejects this account, for it has counter-intuitive consequences. For instance, if *F* and *G* are absolutely structural properties, then the proposed analysis entails that the disjunctive property *being F or G* is equally structural as the conjunctive property *being F and G*. But since disjunctions are less explanatory than conjunctions, and the structural is supposed to be explanatory, disjunctions should be less structural than conjunctions, even though conjunction and disjunction contribute the same amount to the length of metaphysical definitions. Sider never provides us with a replacement account of this flawed definition of comparative structuralness.

Reference is a relation between terms in our language and entities and structures in the world. As we have seen, it is one among many such *reference-like relations*, i.e., relations between terms in our language and entities in the world that preserve the truth of our theory, but assign as the “meanings” of those terms something other than their intended meanings (their referents).^{64,65} Sider’s official view is that reference is the unique maximally structural reference-like relation. Alternatively, there might have been a tie, i.e., multiple reference-like relations that are equally structural. I will argue that both possibilities lead to serious problems.⁶⁶

Evaluating these two possibilities is difficult because they both involve comparisons of structuralness between (reference-like) relations and, as we have seen, Sider never commits to a precise account of what comparative structuralness consists in. But as a matter of simple logic, either the structuralness of a non-absolutely structural relation partly supervenes on the structuralness of its relata or it does not. But I think we can quickly dispense with the second option. If the structuralness of a relation does not supervene on the structuralness of its relata at all, nothing prevents arbitrarily structural relations holding among arbitrarily non-structural relata. But, surely, if a relation holds exclusively among arbitrarily non-structural items, e.g., the permuted predicate extensions or arbitrary mereological fusions generated by Putnam-style arguments, then this must make that relation comparatively non-structural. So I think Sider has to admit that the structuralness of non-absolutely structural relations partly supervenes on the structuralness of their relata.

However, this partial supervenience is in tension with the uniqueness of the reference relation. All of these reference-like relations agree on their first relata (words in our language), so, assuming partial supervenience, differences in their structuralness must supervene on differences in

the structuralness of their second relata, the meanings they assign to our words. What is more, on Sider's view, they must fully supervene on these differences. Sider's claim is that reference is more structural than other reference-like relations because it assigns more structural meanings. If he were to claim that reference is more structural than some other reference-like relation, even though they both assign equally structural meanings, it would no longer be clear what he means by the structuralness of a reference-like relation.

If the structuralness of a reference-like relation supervenes on the structuralness of the items it assigns as meanings, we should expect that a slight decrease in structuralness in one item can be compensated for by a corresponding difference in structuralness in another item. But then why can't one reference-like relation assign a slightly less structural item as the meaning of one singular term and a slightly more structural item as the meaning of another singular term, while remaining overall just as structural as reference itself? For instance, this reference-like relation might assign a slightly more structural meaning to the name "Ted Sider" and a correspondingly slightly less structural meaning to "Immanuel Kant". If Sider wants to maintain that there is a unique maximally structural reference-like relation, this is an assumption that stands in need of explanation as much as anything does. On this view, the possibility of reference is highly sensitive to the measure of comparative structuralness over the set of possible reference-like relations. Sider never makes this measure precise, nor gives any explanation of why it has a maximum. So it is worth considering what other options Sider has.

It would be overall more plausible for Sider to admit that there is no unique such maximally structural reference-like relation. "Reference" would then refer indeterminately to a family of equally structural reference-like relations. Sider has a model for terms like this; because there is no uniquely maximal reference candidate, there is no fact of the matter as to which of the equally maximal reference candidates it refers to. In this case, Sider says, it is a "merely verbal dispute" to what it refers.

But this, combined with the view that the structuralness of a relation supervenes on the structuralness of its relata, potentially undermines Sider's whole picture. Recall that we can compensate for a decrease in the structuralness of one assigned meaning with an increase in the structuralness of another. If this is the case, then for one reference-like relation R that assigns absolutely structural items as the meaning of some of our terms, we should expect there would be another reference-like relation R^* that assigns less than absolutely structural items as the meaning of those terms, but slightly more structural items as the meaning of other terms. But if this is correct, then there is no fact of the matter whether we ever refer to the absolutely structural. Some reference candidates for "reference" will

assign absolutely structural items as the meaning of terms in fundamental metaphysics; some will assign somewhat less absolutely structural meanings to those terms, but more (but still not absolutely) structural meanings to other terms. But since metaphysics (on Sider's view) is concerned with absolute structure, metaphysics will not be possible. It is not determinately the case that we ever succeed in referring to the absolute structure of reality. But this would mean Sider had not explained the possibility of metaphysics. It would remain a "verbal dispute" as to whether we have ever succeeded in talking about the fundamental metaphysical structure of reality.

I have posed a dilemma for Sider: either reference is the uniquely maximally structural reference-like relation, or there is no such unique maximally structural relation. On the first horn, metaphysics is possible, but only if we assume that the metric of comparative structuralness over reference-like relations has a maximum. On this horn, Sider has explained the possibility of metaphysics, but by appeal to an assumption that is as much in need of explanation as anything else. On the second horn, for all Sider has shown, among the family of equally structural reference-like relations, there may be some that never assign absolutely structural meanings. In this case, Sider will have failed to explain the possibility of talk about the absolute structure of reality, i.e., metaphysics.

* * *

Contemporary abductivists thus fail to explain A1–A3.⁶⁷ From a Kantian point of view, therefore, abductivism in metaphysics appears to be a form of dogmatism: 'the prejudice that without criticism reason can make progress in metaphysics' (Bxxx). The abductivist offered the beginnings of an explanation of how metaphysics is possible, through abduction, but did not carry that through by offering an explanation of why abductive inference in metaphysics is knowledge-generating. Abductivism is not wholly uncritical (e.g., it admits that when inference to the best explanation gives out, so too does our justification for metaphysical theorising), but nor does it rise to level of a 'critique of the faculty of reason in general' (Axii). In particular, abductivism dogmatically assumes without explanation that the canons of abductive reasoning that generate knowledge in natural science will do so in metaphysics as well. What is more, our discussion up to this point shows that the charge of "dogmatism" is not simply a Kantian pejorative; it articulates an immanent critique of abductivism. Abductivism, doing metaphysics by inference to the best explanation, does not issue in a good explanation of the possibility of metaphysics. By abductivism's *own lights*, then, a metaphysics that could explain its own possibility would be *ceteris paribus* superior to one that does not. Consequently, such a theory is *ceteris paribus* preferable to abductivism.

15.5 Conclusion

Some abductivists will seek to defend abductive metaphysics through a “good company” objection: any reason to be sceptical about abduction in metaphysics is a reason to be, implausibly, sceptical about abduction in natural science. But this misrepresents the dialectic. Kant accepts that abduction is alethically and explanatorily successful, as long as it restricts itself to the spatiotemporal world and obeys certain regulative (and transcendental) principles (see [Section 15.2](#)). The greater generality of metaphysics leaves logical room for abduction to be successful in natural science without it being successful in metaphysics. Furthermore, the empirical success of natural science over the past several hundred years makes it overwhelmingly plausible that the natural sciences are at least alethically, and perhaps even explanatorily, successful. Metaphysics has no such string of successes to point to, for it is a battlefield of endless controversies. In the face of this, the lack of a satisfying explanation of its possibility is a powerful challenge to the science of metaphysics.

Some will respond that the fact that abductive metaphysics cannot explain its own possibility, i.e., it remains dogmatic, is a compelling reason to reject it only if some other conception of metaphysics can do better. Kant would agree; he does not think that the third option, to reject metaphysics altogether (to “feign indifference” to it), is a viable option for human reason, for our moral vocation (the Highest Good) depends on certain metaphysical foundations (i.e., free will, the existence of God, the immortality of the soul). Kant also, of course, thinks he has an explanation of the possibility of metaphysics, one that relies on transcendental idealism. Whether that explanation succeeds or not, however, is outside the scope of this chapter. In conclusion, I just want to note that even if Kant’s transcendental idealist explanation of metaphysics is beset with insuperable problems, I think, it is, this does not automatically rebound to the credit of the abductivist, for there is a whole separate family of metaphysical views still to be considered: the metaphysical views of the post-Kantian idealists, who agreed with Kant in rejecting pre-Kantian dogmatism and would have agreed with him in rejecting contemporary abductivism, but who thought we had to go beyond Kantian transcendental idealism to obtain a properly critical metaphysics, i.e., one that can explain its own possibility. But that is a story for another time.⁶⁸

Abbreviations for the Work of Kant

All works of Kant are cited according to volume and page number in the ‘Akademie Ausgabe’ (AA): Immanuel Kant. 1900-. *Gesammelte Schriften*. Ed. Vol. 1–22: Preussische Akademie der Wissenschaften, Vol. 23: Deutsche

Akademie der Wissenschaften zu Berlin; from Vol. 24: Akademie der Wissenschaften zu Göttingen. De Gruyter.

A/B	<i>Kritik der reinen Vernunft</i> : A edition (1781, AA 4), B edition (1787, AA 3)
BDG	<i>Der einzig mögliche Beweisgrund zu einer Demonstration des Daseins Gottes</i> (AA 02)
EEKU	Erste Einleitung in die Kritik der Urteilskraft (AA 20)
KU	<i>Kritik der Urteilskraft</i> (AA 5)
Log	Jäsche Logik (AA 9)
MAN	<i>Metaphysische Anfangsgründe der Naturwissenschaften</i> (AA 4)
Refl.	Reflexionen (AA 14–19)
ÜE	<i>Über eine Entdeckung, nach der alle neue Kritik der reinen Vernunft durch eine entbehrlich gemacht werden soll</i> (AA 8)
V-Lo/Blomberg	Logik Blomberg (AA 24)
V-Lo/Busolt	Logik Busolt (AA 24)
Vo-Lo/Dohna	Logik Dohna-Wundlacken (AA 24)
V-Lo/Philippi	Logik Philippi (AA 24)
V-Lo/Pölitz	Logik Pölitz (AA 24)
V-Lo/Wiener	Wiener Logik (AA 24)
V-Met/Dohna	Metaphysik Dohna (AA 28)
V-Met/Herder	Metaphysik Herder (AA 28)
V-Met/Mron	Metaphysik Mrongovius (AA 29)
V-Met-L ₂ /Pölitz	Metaphysik Pölitz (AA 28)
V-Met/Schön	Metaphysik von Schön (AA 28)
V-Met/Volckmann	Metaphysik Volckmann (AA 28)

Notes

- 1 Here, and henceforth, “sciences” refers simply to natural sciences.
- 2 I am going to assume that it is knowledge that the abductivist is after in metaphysics. Whether the abductivist could set their sights on a lower epistemic status (e.g., justified true belief) I will not consider here.
- 3 “Abductivism” refers to a set of doctrines shared by many contemporary metaphysicians, not to a single thesis. So I am not claiming that *scientific realism* follows logically from *abductive methodology* or *abductive epistemology* (it doesn’t). I am claiming only that abductivism, as actually endorsed by practising metaphysicians, contains this commitment.
- 4 A369, A491/B519. In Kantian terms, the primary topic of contemporary metaphysics is things in themselves, although there is also some contemporary work on the metaphysics of appearance as well (e.g., McDaniel 2017: 140–169).
- 5 For example, Biggs 2011; Hawley 2006; Paul 2012; Sider 2009: 385, 2011: 12–15; and Williamson 2016. Some of these references are drawn from Biggs and Wilson 2017: 739.

- 6 Technically, for Kant, a critique is a critique of a *capacity* (Axii), e.g., practical reason or the power of judgement. Without too much distortion, though, we can see abduction as capacity as well, although not all abductivists would agree that it is a basic capacity (as argued in Biggs and Wilson 2017).
- 7 This does not mean it is a successful critique, i.e., a fully adequate explanation of how metaphysics is possible. My argument in this chapter is precisely that it does not succeed in that ambition.
- 8 All abductive metaphysicians with whom I am familiar admit that metaphysics is less abductively secure than natural science, and that there are questions in metaphysics that are epistemically undecidable because there are not sufficiently good grounds to abduct to a single answer or set of answers, though they differ on where the “limits” of metaphysics lie (cf. Bennett 2009 and Sider 2011: 12).
- 9 For Kant, metaphysics is concerned with conditions (grounds), either the conditions of possible experience (immanent metaphysics) or the conditions of supersensible objects (transcendent metaphysics). See A845/B873.
- 10 For Kant there is also transcendent metaphysics of the mind-independent (supersensible), but that is either grounded practically (in the moral law) or is merely speculative and fails to constitute cognition. In either case, it lies outside the scope of this chapter.
- 11 Log 9: 84–85; V-Lo/Dohna 24: 743; V-Lo/Wiener 24: 888.
- 12 Axv; Log 9: 86; Refl. 2680, 16: 467; Refl. 5645, 18: 293; V-Lo/Busolt 24: 647; V-Lo/Pölitz 24: 559. Cf. Bxxii on how what is initially merely the Copernican hypothesis (that objects conform to our capacity for cognition) becomes apodictically certain through the *KrV* itself.
- 13 There are two other Kantian reasons to reject abductive metaphysics. First, metaphysics must be *a priori*, but abduction is empirical. However, few contemporary metaphysicians would accept that metaphysics must be *a priori*, and some even deny that abduction is empirical (e.g., Biggs and Wilson 2017). Second, the abductive conclusion functions initially as a hypothesis, and the hypothesis must be known apodictically to be at least really possible (A770/B798). But this just reiterates the questionable assumption that metaphysics must be apodictic at the level of the possibility.
- 14 *MAN*, 4: 468.
- 15 As Kant argues in the Appendix to the Transcendental Dialectic, “On the regulative use of ideas of pure reason” (A642–668/B670–696). See Section 15.4 for discussion.
- 16 For example, Leibniz gives an essentially abductive justification for pre-established harmony in the *New System of Nature*; see Leibniz 1989: 144–145.
- 17 In *BDG*, Kant endorses an essentially abductive argument for the existence of a ‘rational Author’ from the ‘purposeful provisions’ in all things, in addition to the more famous apodictic proof from the real possibility of all things (2: 159).
- 18 Locke is a possible exception, but further exploring that connection lies outside the scope of this chapter.
- 19 The *locus classicus* is Lipton 2004.
- 20 I take it that this distinction should be acceptable to both those whom Raven (2015) calls “unionists” about metaphysical explanation (e.g., Dasgupta 2014, Fine 2012, and Rosen 2010), and those he calls “separatists” (e.g., Audi 2012, Schaffer 2012).
- 21 One can have a more deflationary conception of explanation (e.g., a broadly pragmatist one) but this will push one’s metaphysics in a more deflationary direction. Cf. Dasgupta 2018.

- 22 I have put this point at the level of language, but it could also be put at the level of thought: our concepts must carve the world at its explanatory joints.
- 23 Assuming grounding is a relation between facts. If not, adjust the example accordingly.
- 24 See the texts cited in the Introduction.
- 25 This material is presented more fully in Stang n.d.1.
- 26 I include the qualification ‘in virtue of its very form’ because how, and whether, the matter of experience grounds the matter of the objects of experience is complex and controversial. For discussion, see Stang 2016a and 2018.
- 27 MAN 4: 468, 469. Kant there talks of nature, but in the case of causally efficacious substances (which is what laws are about), natures are real essences (see 4: 468n.).
- 28 For further discussion, see Stang 2016b: 234–244.
- 29 This raises the question of which is explanatorily prior: law or real essence. I think real essence is prior; the necessity of laws is grounded in real essences. I argue for this at greater length in Stang 2016b.
- 30 ÜE 8: 238, Log 9: 143, Br. 11: 37, V-Lo/Wiener 24: 839, and V-Lo/Dohna 24: 757.
- 31 Log 9: 61, ÜE 8: 229, Br. 11: 36, Refl. 5706, V-Lo/Blomberg 24: 116, V-Lo/Philippi 24: 408, 456, V-Lo/Pölit, 535–536, V-Lo/Busolt 24: 634, V-Lo/Dohna 24: 727, V-Lo/Wiener 24: 839, V-Met-L₂/Pölit, 28: 553.
- 32 MAN, 4: 509, 511; V-Met-L₂/Pölit, 28: 553.
- 33 A727–728/B755–756, Refl. 3966, V-Lo/Blomberg 24: 116–117, V-Lo/Dohna 24: 757, V-Lo/Wiener 24: 919.
- 34 This is only true of *given* empirical concepts; in the case of *made* empirical concepts (concepts of artefacts, concepts of invented fictions), the logical essence and the real essence coincide. For the given-made distinction, see Log 9: 93.
- 35 “Essentialist” readings of Kantian laws of nature similar to that presented here and in Stang 2016b have also been defended by Watkins (2005) and Kreines (2008).
- 36 V-Lo/Volckmann 28: 422; V-Met/Schön 28: 503; V-Met-L₂/Pölit 28: 560.
- 37 This of course just raises the question of the ontological statuses of essences themselves. Are they also part of Kant’s “ontology”? For reasons of space, I will not address that question here.
- 38 I want to remain neutral on whether the identity of the natural kind *fully* individuates empirical concepts, i.e., whether, as seems plausible, there could be two distinct empirical concepts of one and the same natural kind (e.g., $\langle \text{water} \rangle$ and $\langle H_2O \rangle$).
- 39 Kant thinks we only ever know, at most, parts of the real essences of things; see Br 11: 37, V-Lo/Dohna 24: 728, V-Met-L₂/Pölit 28: 553.
- 40 This is why Kant claims that not merely the possession, but the very possibility, of an empirical concept of a force depends upon the actual existence of that force in space and time. If that force did not exist (if a force without its real essence did not exist), no concept we form would be a concept of that very force. See B252, MAN 4: 486–487.
- 41 This follows from a point Kant repeatedly makes in his lectures – that the marks contained in (given) empirical concepts are typically so impoverished as to be virtually uninformative; see LB, 24: 116, 117–118, 271; WL 24: 919, A727–728/B755–756.
- 42 By “there being” a set of empirical concepts I mean that were we to engage in the relevant acts of concept formation, we would form those concepts. So the “being” of those concepts, for the purposes of this argument, is independent of whichever acts of concept formation we actually perform.

- 43 A368; V-Met/Volckmann, 28: 401, 408; V-Met-L2/Pölitz 28: 548; V-Met/Dohna 28: 624, 628; V-Met/Mron 29: 808, 818.
- 44 They are impure because they involve the empirical concept of matter (see B for the notion of the impure *a priori*). For the transcendental-metaphysical distinction, see *MFNS* 4: 469–470; *KU* 5: 181.
- 45 In the terms of the Introduction, it produces only comparative, not apodictic, certainty.
- 46 See A650–651/B678–679, where a transcendental principle is said to underlie the logical/regulative principle of systematicity as such; at A656/B684 he claims there is a transcendental principle of specification and, at A660/B688, of continuity. Pickering (2011) reads these references to a transcendental principle, implausibly, as references to transcendental illusion.
- 47 I agree with Geiger (2003: 275–282) that the systematicity of nature is a transcendental principle (and with his trenchant criticisms of prior readings), but I do not accept his argument for that conclusion.
- 48 This is because “experience” for Kant has a much richer, more determinate, and more scientific content than the relatively minimal “experience” of the empiricist tradition. See Stang 2018 for discussion.
- 49 A slightly stronger conclusion is warranted: since every object is governed by causal laws, and laws concern natural kinds (objects in virtue of a common real essence), every object belongs to some common natural kind.
- 50 An especially clear statement of the problem occurs at *EEKU* 20: 208–209; see also *KU* 5: 183, 186, 188.
- 51 *KU* 5: 180.
- 52 It might seem obvious to some readers that Kant’s explanation can only be an epistemic one (of our rational warrant for assuming K2) rather than a metaphysical one (of the truth of K2). But, arguably, this depends upon assuming a transcendental realist view of Nature on which the “being” of Nature can transcend what we, in principle, could have rational warrant for judging about it (i.e., assuming that the epistemology and metaphysics of Nature could come that far apart). But I will leave this issue aside.
- 53 In conversations with abductive metaphysicians, I have found this a common manoeuvre.
- 54 A more strictly naturalist position, on which metaphysics is simply identical to natural science, is not targeted by the arguments of this chapter.
- 55 It is even clearer that Kant himself cannot accept this answer, for natural scientific abduction is limited to objects of possible experience, while metaphysics makes claims that are either beyond the bounds of possible experience (transcendent metaphysics) or claims about the forms of experience themselves (immanent metaphysics). That abduction in the former cause is successful is no reason to think it would be successful in the latter cases. For the immanent-transcendent distinction, see A846/B874.
- 56 Schaffer (2003) raises the intriguing possibility of an infinite downward regress of (in my terms) ontic explanatory structure, but I will not pursue that thought any further here.
- 57 Notice that a Bayesian account of abduction alone will not suffice: even if we build our canons of abductive theory choice into our priors (e.g., assigning a prior probability to more unified explanations), this will not account for why reasoning according to such priors constitutes knowledge (why it non-accidentally tracks the truth). What is more, those priors must be formulated in a language that carves the world at its explanatory joints. So a Bayesian account by itself will not

- explain A2 or A3. See Bradley 2020 for further discussion. Thanks to Trevor Teitel for helpful conversations about this issue.
- 58 Williams 2019 offers an interpretationist meta-semantics in a similar spirit, although it departs from Sider in key respects (56: n. 26; 67: n. 10). Unlike Sider, though, Williams does not explicitly apply it to reference in metaphysics. In Stang n.d.2, I argue that Williams's theory does not adequately explain why metaphysics is semantically possible.
- 59 I follow Sider in making the simplifying assumption that our theory is first-order. This assumption is potentially fraught, however, because, as Bayes (2001) shows, Putnam's original argument fails in higher-order theories. But I will not pursue this issue further here.
- 60 The same idea underlies Quine's argument for the inscrutability of reference; see Quine 1960. Putnam's original argument used the upwards and downwards Löwenheim-Skolem theorems to show that if our theory has any infinite models, for any infinite cardinality it has a model of that cardinality. Lewis (1984) points out that the main result follows from much simpler considerations about the permutation of the domain.
- 61 In fact, it is more complicated than this; see below.
- 62 Sider 2011: 128. Because anything of any syntactic type whatsoever can be evaluated for its structuralness, I choose the dummy term "item". Bear in mind that the structuralness of a term in our language is, in general, distinct from the structuralness of its referent. Even if spacetime points are absolutely structural, names for them are not, for names are not (according to Sider) part of the fundamental structure of reality.
- 63 Sider 2011: 129–133.
- 64 Following Frege, it is standard in analytic philosophy to distinguish between two notions of meaning, sense and reference. In line with Sider's approach, I am ignoring sense and focusing only on meaning as reference.
- 65 I will speak of reference-like relations as assigning "meanings" to terms, because it would be needlessly confusing to speak of them as assigning "referents" to them (only reference assigns their referents) and needlessly prolix to talk about "their semantic value assigned them by that reference-like relation" (or something of that ilk). But just to be clear: the meaning of some term according to a reference-like relation is, in general, not its intended or actual meaning (its referent), unless that reference-like relation is reference itself.
- 66 Talk about reference-like relations is just a way of talking about interpretations of the language different from the intended interpretation. But for ease of exposition I focus on the (reference-like) relation between terms and the meanings assigned to it by different interpretations, rather than on the interpretations themselves.
- 67 This does not mean that abductivism cannot explain A1–A3. However, I think my discussion so far shows that the prospects for such an abductivist explanation are bleak.
- 68 I would like to thank audiences at the University of Kansas and the participants in a "block seminar" at the University of Bonn for their helpful feedback on earlier presentations of this material. Brad Cokelet, Catharine Diehl, Alex Englander, Tyler Hildebrandt, Karl Schafer, and Trevor Teitel gave me invaluable feedback and comments as well. Special thanks go to Robb Dunphy for extensive comments on the penultimate draft, which vastly improved the (still imperfect) final chapter.

Bibliography

- Audi, P. (2012) "Grounding: Toward a Theory of the *In-Virtue-Of* Relation" *Journal of Philosophy* Volume 109, pp. 685–711
- Bayes, T. (2001) "On Putnam and His Models" *Journal of Philosophy* Volume 98(7), pp. 331–350
- Bennett, K. (2009) "Composition, Colocation, and Metaontology". In Chalmers, D., Manley, D., and Wasserman, R. (Eds.) (2009) *Metametaphysics: New Essays on the Foundations of Ontology* (Oxford: Oxford University Press)
- Biggs, S. (2011) "Abduction and Modality" *Philosophy and Phenomenological Research* Volume 83, pp. 283–326
- Biggs, S. and Wilson, J. (2017) "The A Priority of Abduction" *Philosophical Studies* Volume 174, pp. 735–758
- Bradley, D. (2020) "Naturalness as a Constraint on Priors" *Mind* Volume 129(513), pp. 179–203
- Chalmers, D., Manley, D., and Wasserman, R. (Eds.) (2009) *Metametaphysics: New Essays on the Foundations of Ontology* (Oxford: Oxford University Press)
- Correia, F. and Schnieder, B. (2012) *Metaphysical Grounding: Understanding the Structure of Reality* (Cambridge: Cambridge University Press)
- Dasgupta, S. (2014) "On the Plurality of Grounds" *Philosophers' Imprint* Volume 14, pp.1–28
- Dasgupta, S. (2018) "Realism and the Absence of Value" *Philosophical Review* Volume 127(3), pp. 279–322
- Fine, K. (2012) "Guide to Ground". In Correia, F. and Schnieder, B. (2012) *Metaphysical Grounding: Understanding the Structure of Reality* (Cambridge: Cambridge University Press).
- Geiger, I. (2003) "Is the Assumption of a Systematic Whole of Empirical Concepts a Necessary Condition of Knowledge?" *Kant Studien* Volume 94(3), pp. 273–298
- Goodman, N. (1944) "The New Riddle of Induction". In *Fact, Fiction, and Forecast* (Cambridge, MA: Harvard University Press)
- Hawley, K. (2006) "Science as a Guide to Metaphysics?" *Synthese* Volume 149, pp. 451–470
- Kreines, J. (2008) "Kant on the Laws of Nature: Laws, Necessitation, and the Limitation of Our Knowledge" *European Journal of Philosophy* Volume 17(4), pp. 527–558
- Leibniz, G.W. (1989) *Philosophical Essays* R. Ariew and D. Garber (Trans. and Eds.) (Hackett)
- Lewis, D. (1984) "Putnam's Paradox" *Australasian Journal of Philosophy* Volume 62(3), pp. 221–236
- Lipton, P. (2004) *Inference to the Best Explanation*, 2nd ed. (London: Routledge)
- McDaniel, K. (2017) *The Fragmentation of Being* (Oxford: Oxford University Press)
- Paul, L. (2012) "Metaphysics as Modeling: The Handmaid's Tale" *Philosophical Studies* Volume 160, pp. 1–29
- Pickering, M. (2011) "The Idea of the Systematic Unity of Nature as a Transcendental Illusion" *Kantian Review* Volume 16(3), pp. 429–448

- Putnam, H. (1981) *Reason, Truth, and History* (Cambridge: Cambridge University Press)
- Quine, W.V.O. (1960) *Word and Object* (Cambridge, MA: Harvard University Press)
- Raven, M. (2015) "Ground" *Philosophers' Compass* Volume 10(5), pp. 322–333
- Rosen, G. (2010) "Metaphysical Dependence: Grounding and Reduction". In B. Hale and A. Hoffmann (Eds.) *Modality: Metaphysics, Logic, and Epistemology* (Oxford: Oxford University Press)
- Schaffer J. (2003) "Is There a Fundamental Level?" *Noûs* Volume 37(3), pp. 498–517
- Schaffer, J. (2009) "On What Grounds What". In Chalmers, D., Manley, D., and Wasserman, R. (Eds.) (2009) *Metametaphysics: New Essays on the Foundations of Ontology* (Oxford: Oxford University Press)
- Schaffer, J. (2012) "Grounding, Transitivity, and Contrastivity". In Correia, F. and Schnieder, B. (2012) *Metaphysical Grounding: Understanding the Structure of Reality* (Cambridge: Cambridge University Press)
- Sider, T. (2009) "Ontological antirealism". In Chalmers, D., Manley, D., and Wasserman, R. (Eds.) (2009) *Metametaphysics: New Essays on the Foundations of Ontology* (Oxford: Oxford University Press)
- Sider, T. (2011) *Writing the Book of the World* (Oxford: Oxford University Press)
- Stang, N. (2015) "Who's Afraid of Double Affection?" *Philosophers' Imprint* Volume 15(18), pp. 1–28
- Stang, N. (2016a) "Appearances and Things in Themselves: Actuality and Identity" *Kantian Review* Volume 21(2), pp. 283–292
- Stang, N. (2016b) *Kant's Modal Metaphysics* (Oxford: Oxford University Press)
- Stang, N. (2018) "Kant, Cohen, and the Concept of Experience". In C. Damböck (Ed.) *Philosophie und Wissenschaft bei Hermann Cohen* (Cham: Springer), pp. 13–40
- Stang, N. (n.d.1) "Kant on Natural Kinds". Unpublished manuscript
- Stang, N. (n.d.2) "Is Kant's Critique of Metaphysics Obsolete?" <https://philpapers.org/rec/STAIKC>
- Watkins, E. (2005) *Kant and the Metaphysics of Causality* (Cambridge: Cambridge University Press)
- Williams, J.R.G. (2019) *The Metaphysics of Intentionality* (Oxford: Oxford University Press)
- Williamson, T. (2016) "Abductive Philosophy" *The Philosophical Forum* Volume 47(3–4), pp. 263–280