Introduction

This book is about Immanuel Kant's account of reason as the source of metaphysical speculation, as he develops it in the Transcendental Dialectic of the *Critique of Pure Reason*. It has two major aims. First, it will offer a novel interpretation of the Transcendental Dialectic that isolates its constructive side (Kant's account of the rational sources of speculative metaphysics, concerning the soul, the world as a whole, and God), and that distinguishes it clearly from its destructive side (Kant's *critique* of this kind of metaphysics). We will see that Kant himself does not always keep these two projects separate, with the result that there are passages in the Transcendental Dialectic that appear puzzling or even confused but that make perfect sense once the two strands are clearly distinguished. I will offer interpretations of all main parts of the Transcendental Dialectic, and even though I cannot claim to be able to solve all the exegetical problems Kant's texts present (some of his arguments and derivations are extremely brief and cryptic), the reading proposed here unveils a consistent and philosophically attractive account of metaphysical thinking that has so far been widely ignored in the literature on Kant.

Second, this book will reconstruct, and where possible defend, a Kantian account of the rational sources of metaphysical thinking. In particular, it will argue that Kant is right in claiming that metaphysical speculation arises naturally out of principles that guide us in everyday rational thought. On the one hand, the structure of rational thinking is *discursive* and *iterative*, requiring us to ask not only for explanations of empirical phenomena but also for explanations of the phenomena we rely on to explain them (and for their explanations, and so on). On the other hand, as rational inquirers we want our questions to come to a satisfactory conclusion, which they can find only in *ultimate* answers, that is, in answers that do not raise further questions of the same kind. As I will argue, Kant gives us good reason to think that *discursivity*, *iteration*, and striving for *completeness* are fundamental features of rational thinking and that, taken together, they give rise to a specific kind of metaphysical speculation. This is a distinctive and original perspective on metaphysics that deserves to be taken seriously in the current metaphysical and metametaphysical debates. As I will indicate in passing in the course of this book, many

of the metaphysical issues Kant discusses in the Transcendental Dialectic are still very much alive today.

0.1 Beginning at the Beginning(s)

Kant begins the *Critique of Pure Reason*, his treatise on method in metaphysics (Bxxii), with an implicit reference to the beginning of the first book ever to bear that title. Aristotle famously starts his *Metaphysics* with the observation that human beings by their nature desire to know (*eidenai*) (*Metaphysics* 980a1) and then explains that knowledge starts from sense experience and ends with knowledge of first principles and causes. Aristotle calls the science that investigates these principles and causes 'first philosophy'; later, it will be known as 'metaphysics.' Hence, we can sum up Aristotle's line of thought by saying that human beings, because of their rational nature, strive for knowledge, beginning with knowledge from experience and ending with metaphysical knowledge of first principles. Now compare this with how Kant begins the Preface of the first edition of his *Critique of Pure Reason*:

> Human reason has the peculiar fate in one species of its cognitions that it is burdened with questions which it cannot dismiss, since they are given to it as problems by the nature of reason itself, but which it also cannot answer, since they transcend every capacity of human reason. Reason falls into this perplexity through no fault of its own. It begins from principles whose use is unavoidable in the course of experience . . . With these principles it rises (as its nature also requires) ever higher, to more remote conditions . . . [R]eason sees itself necessitated to take refuge in principles that overstep all possible use in experience . . . [I]t thereby falls into obscurity and contradictions . . . The battlefield of these endless controversies is called *metaphysics*. (Avii–viii)

On the most fundamental point, Kant agrees with Aristotle: it is part of human nature – Kant speaks of the nature of human reason, but also of human nature (Ax) – to strive for metaphysical cognition and knowledge.[1] This is knowledge of first principles, according to Aristotle, and cognition of higher and highest principles, according to Kant. Kant also agrees, at least in broad outline, that sense experience is the basis for all other knowledge by admitting that reason starts from principles used in experience. As he writes some 300 pages later in the first *Critique*, "[a]ll our cognition starts from the senses, goes from there to the understanding, and ends with reason" (A298f/B355). But where Aristotle presumably thinks that, by rising from experience to reason, metaphysical knowledge is in fact to be gained, Kant, after two thousand years of metaphysical speculation, can only make out a battlefield scattered with the ruins of

[1] In this book, the term 'cognition,' if used without further qualification, means 'theoretical cognition.' Note that for Kant, cognition is not the same as knowledge (Willaschek and Watkins 2017 and Chapter 9, note 12 below).

failed theories. According to Kant, human reason, by rising to ever higher principles and more remote conditions, overreaches itself and falls into fallacies and contradictions. The fate of human reason is thus a truly tragic one. The tendency not just to ask metaphysical questions but also to devise answers to them is built into the very structure of rational thought. At the same time, rational thought is limited in ways that make it impossible for us ever to know which answers to these metaphysical questions are correct and that make it seem doubtful that these questions make sense in the first place.

0.2 The Rational Sources Account

So why does Kant hold that human reason inevitably confronts us with metaphysical questions? In the Introduction to the second edition (the 'B-edition') of the *Critique of Pure Reason*, Kant distinguishes between metaphysics "as a science" (which we do not yet possess) and "metaphysics as a natural predisposition" (B22), that is, metaphysics as a natural tendency in human beings to ask metaphysical questions and to come up with answers to them. Kant asks:

> How is metaphysics as a natural predisposition possible? i.e. how do the questions that pure reason raises, and which it is driven by its own need to answer as well as it can, arise from the nature of universal human reason? (B22)

There are three claims implicit in this question: (1) pure reason – rational thought independently of any input from the senses – raises metaphysical questions; (2) pure reason is driven by its own need or its nature to answer these questions, even if the answers may not be ultimately warranted ('as well as it can'); and (3) the metaphysical questions do not arise from the minds of individual and perhaps misguided metaphysicians but rather have their source in 'the nature of universal human reason' – that is, they arise from the very structure of rational thinking as such. I will call the conjunction of these three claims 'Kant's account of the rational sources of metaphysics,' or the *Rational Sources Account* for short. The Transcendental Dialectic of the first *Critique* contains Kant's extended argument for this account, with the general framework being developed in the Introduction and Book 1 of the Dialectic and the specifics filled in in the chapters on the paralogisms, the antinomies, and the ideal of pure reason and in the Appendix.

An important aspect of the Rational Sources Account concerns the relation between reason in general and 'pure' reason. According to Kant, even though it is *pure reason* that raises and attempts to answer metaphysical questions, it is reason in general, or *universal human reason*, from which these questions originally arise. This means that even though metaphysical questions take us beyond the bounds of possible experience, they are not willful speculations but

arise naturally from features of rational thinking that are at work even in the most ordinary empirical employments of reason in everyday life. It is this latter claim that makes Kant's diagnosis of metaphysical thinking philosophically deep and attractive. Although certain modes of metaphysical thinking may be fundamentally flawed and may have to be abandoned (or, as Kant argues, transformed into a practical mode), they have their ultimate source in something that we cannot possibly abandon, namely, rational thinking. As Kant explains, it is the task of a "critique of pure reason" to determine "the sources, as well as the extent and boundaries" of metaphysics (Axii). His central result in this respect is that the *sources of metaphysics* lie in *reason* itself (A309/B366).

In its broadest outline, this is how the story goes: we begin with principles that work fine within the realm of experience, for instance, the principle that every alteration must have a cause, or that in every change there must be something that persists. "With these principles," Kant says in the A-Preface, reason "rises (as its nature also requires) ever higher, to more remote conditions" (Avii), for example, by asking what caused the *cause* of the original alteration to occur or by wondering whether the very thing that persisted in one situation cannot be transformed into something else in a different situation, thus in turn requiring something that persists. Only much later in his book, in the Introduction to the Transcendental Dialectic, does Kant explain why it lies in the nature of reason to ask these kinds of questions. There, Kant argues that whenever we are confronted with something 'conditioned' (roughly, something calling for an explanation), reason compels us to look for its 'condition' (something that explains it) – a tendency Kant traces back to a core function of human reason: syllogistic reasoning. Just as we can seek premises for given conclusions, and then further premises from which to derive the original premises, we ask not just for conditions of the conditioned but also for the conditions of the conditions, etc. In this way, we find ourselves starting on a regress that is potentially infinite. As Kant says of reason in the A-Preface: "its business must always remain incomplete because the questions never cease" (Aviii). The questions can come to an end only in something 'unconditioned,' for instance in an uncaused cause or a substance that persists throughout every possible change.

In this way, asking for explanations of ordinary phenomena ultimately leads us to assume the existence of something 'unconditioned.' If there is a fire, for instance, we ask what caused it. If lightning caused the fire, we ask what caused the lightning. If electrical charges in the air caused the lightning, we ask what caused the charges, etc.: 'the questions never cease.' But if we assume that there is an answer to each and every one of these causal questions, it seems that there must either be an uncaused cause – a cause that does not raise a further question about its cause – or a series of causes and

effects that stretches infinitely back in time. In both cases, Kant argues, we have posited something unconditioned, since neither a first member in a series of causes nor a complete infinite series of causes can itself have a cause that belongs to the same series.

Since everything we experience is conditioned in one way or another (spatially, temporally, causally, etc.), in searching for something *un*conditioned we leave the field of experience and enter the realm of metaphysical speculation: "For that which necessarily drives us to go beyond the boundaries of experience and all appearances is the *unconditioned*, which reason necessarily and with every right demands in things in themselves for everything that is conditioned, thereby demanding the series of conditioned as something completed" (Bxx).[2]

So here we have the outline of an explanation of why metaphysical questions arise from the very nature of rational thought. Rational thinking includes a tendency to move from the conditioned to its condition, a movement innocently at work in syllogistic reasoning and empirical explanation, but one that goes overboard when it aspires to completeness (finding all conditions, giving ultimate answers) because it then leads us to metaphysical claims about uncaused causes, absolute substances, and the like. The concept of the unconditioned that we employ in these metaphysical speculations is not abstracted from experience but based on the reliable activity of syllogistic reasoning and hence comes with the best rational credentials. As Kant argues extensively in the Transcendental Aesthetic and Transcendental Analytic of the first *Critique*, however, human cognition is limited to the realm of possible experience, which implies that we can have no cognition of the unconditioned. Hence, the metaphysical theories that grow naturally out of rational thinking and that seem to provide us with metaphysical cognition and knowledge, according to Kant, lead us into fallacies and contradiction.

Kant's claim that there is a natural disposition toward metaphysics can thus be stated more explicitly as follows:

RS-1 Rational reflection on empirical questions necessarily raises *metaphysical questions* about 'the unconditioned.'

RS-2 Rational reflection (by 'pure reason') on these metaphysical questions necessarily leads to *metaphysical answers* that appear to be rationally warranted.

RS-3 The rational principles that lead from empirical to metaphysical questions and from there to metaphysical answers are principles of '*universal human reason*'; that is, they belong to rational thinking as such.

[2] Why Kant restricts this claim to things in themselves will concern us later (5.3). We can set this question aside for now.

The central *philosophical* thesis of this book will be that Kant indeed discovered a source of metaphysical thinking that lies in reason itself. Reason, according to Kant, is characterized by three features that, taken together, lead to metaphysical speculation. First, there is the *discursive* character of human thinking. Human reason, according to Kant, is not intuitive but discursive in that its cognitions result from actively and sequentially processing a multitude of elements. (Contrast sensible intuition, which is passive and holistic.) The need to ask questions and to require grounds and explanations is an expression of this kind of discursivity since it involves the distinctions between question and answer, grounded and ground, *explanandum* and *explanans*, and the active transition from the one to the other. Second, reason-giving and rational explanation are *iterative*. If 'Why A?' is a good question and 'Because of B' is a good answer, then 'Why B?' is a good question too – one that needs to be answered if reason is to be satisfied. And third, there is the rational need for *completeness* – for complete explanations and ultimate answers. As rational inquirers, we cannot be wholly satisfied until we arrive at an answer that does not raise further questions (of the same kind). It is the combination of these three features that takes us, in Kant's words, from the conditioned, through the complete series of conditions, to the unconditioned.

The Kantian account of the sources of metaphysical speculation differs from earlier critiques of metaphysics, e.g. those from empiricist philosophers such as Bacon and Hume, in that it traces both the metaphysical urge and the failure of metaphysics not to the contingent shortcomings of individual thinkers or to aspects of human psychology but rather to the very structure of rational thinking itself. If true, this is a deep and important insight. Uncovering this insight will require some work, however, since Kant's main treatment of this issue in the Transcendental Dialectic of the *Critique of Pure Reason* is highly complex and often difficult to follow.

0.3 One Argument, Four Levels

In the Transcendental Dialectic, we can distinguish four levels on which Kant's Rational Sources Account operates. These levels roughly correspond to the four main parts of the Transcendental Dialectic (Introduction, Book One, Book Two, Appendix). I speak of 'levels' and not 'steps' of an argument because Kant's unfolding of the Rational Sources Account is not best read as consisting of a series of consecutive steps, each of which is foundational to, and independent of, the next. Rather, I suggest that we understand these levels as parts of a complex argument that first lays out a general framework and then fills in the details as it proceeds.

On the first, most general level, there is the transition from the 'Logical Maxim,' which requires us to find a condition for each conditioned *cognition*,

to the 'supreme principle of pure reason' (or 'Supreme Principle'), according to which if something conditioned is given, then so is the complete series of conditions, where this series itself is unconditioned. On Kant's view, it is this latter principle that drives human reason to metaphysical speculation. Kant motivates the transition from the Logical Maxim to the Supreme Principle in the second part of the Introduction to the Transcendental Dialectic (A305/B362–A309/B366). At the same time, Kant establishes the general framework he then also applies, *mutatis mutandis*, on the following levels, namely, a move from a 'logical use of reason' to its 'real' or 'transcendental' use and (as becomes fully explicit only in the Appendix to the Transcendental Dialectic) from a 'regulative' use of reason's principles and ideas to a 'constitutive' use of them (more on which soon).

On the second level, Kant derives the *system of transcendental ideas*. Transcendental ideas are concepts of objects that, if they do exist, are unconditioned (such as the soul, the world, and God). According to Kant, these ideas arise naturally out of 'necessary inferences of reason' and can be brought into a system that guarantees the completeness of our account of them. This system consists of the concept of the unconditioned (which is the '*common title* of all ideas of reason'), the three *classes* of transcendental ideas (psychological, cosmological, and theological, corresponding to the ideas of soul, world, and God), and nine *modes* (ways in which objects can be thought to be unconditioned): substantiality, simplicity, personality, and spirituality in the case of the soul (A344/B402; B419); the absolute completeness of composition, division, origin, and mutual dependence in the case of the world (A415/B443); and finally the idea of an *ens realissimum* in the case of God (A571/B599–A583/B611). On this second level, Kant only derives the *systematic order* of transcendental ideas, not these ideas themselves.[3]

On the third level, there are the specific 'dialectical' (that is, fallacious) inferences that purport to provide us with a priori knowledge about the soul, the world, and God (the 'paralogisms,' the arguments leading up to the 'antinomies,' and the (one) argument for the existence of God). These inferences, Kant claims, have their source in human reason itself and must therefore appear compelling to anyone unequipped with the results of Kant's *Critique of Pure Reason*. At the same time, these inferences deliver the specific transcendental ideas that fall into the classes and modes derived on the previous level.

Finally, on the fourth level, Kant argues that even though the transcendental principles and ideas derived at the previous level have a legitimate 'regulative' use in guiding empirical scientific research and our search for unity in the

[3] This is not how Kant's derivation of the transcendental ideas is usually understood; I will develop my reading in detail in Chapter 6.

diversity of natural phenomena, they are easily mistaken for being constitutive, that is, for being true representations of objects. Taken together, Kant's reflections and arguments on these four levels are meant to show that metaphysical speculation about the unconditioned, in its various forms, arises naturally and inevitably out of the very structure of human reason.

According to Jonathan Bennett, Kant's argument for the Rational Sources Account and the conception of reason on which it relies "is a clumsy attempt to rationalize a set of problems which reflect not the structure of reason but the preoccupations of German academic philosophers at the time when Kant was writing. Where the theory has an effect, it is by tempting Kant into a brutal and insensitive forcing of his material into unnatural shapes and never by genuinely illuminating it" (Bennett 1974: 258). And indeed, if read as consecutive argumentative steps, Kant's way of developing the Rational Sources Account will not look very convincing.[4] Against readers like Bennett, however, I will argue that in fact Kant offers a highly complex argument for the Rational Sources Account that rests on an equally complex account of human reason. Uncovering the Rational Sources Account as a distinctive line of thought in the Transcendental Dialectic will be the central *exegetical* result of this book.

Specifically, each of the four levels outlined earlier involves a transition from 'logical' concepts and principles to 'transcendental' ones (or, as Kant also puts it, from the 'logical use of reason' to its 'real use'), where logical principles abstract from the objects of cognition and consider only formal relations between them, while transcendental principles consider cognitions in relation to their objects. For instance, the Logical Maxim mentioned on the first level of Kant's account requires us to search, for each cognition, for a more general cognition from which it follows. This is a 'logical' project in that it abstracts from the content of our cognitions and considers only their logical entailment relations. Kant therefore attributes it to the 'logical use of reason.' By contrast, the Supreme Principle requires us to look for a condition for each conditioned *object*, thus moving from mere logical conditioning relations among *cognitions* to 'real' conditioning relations among *things*. This is part of what Kant calls the 'real use of reason.'

In this respect, Kant's leading idea is that there is a natural tendency, first, to move from logical principles implicit in universal human reason to the transcendental principles of pure reason, and second, to misunderstand these principles as 'constitutive' (as implying true claims about objects), while their only legitimate use is 'regulative' (that is, directing our search for systematic cognition and knowledge). Concerning the first aspect, a transition from the

[4] For instance, it is only on the second level that the general idea of the unconditioned is introduced, even though it is already employed in the formulation of the principles Kant discusses on the first level.

logical to the transcendental can be found on each of the four levels of Kant's argument. It is modeled on the transition from the table of judgments to the categories (the so-called Metaphysical Deduction) in the Transcendental Analytic, thus revealing the deep structural unity of Kant's overall argumentative strategy in the Transcendental Logic.

On its own, however, the transition from the logical to the real or transcendental use of reason only allows us to ask metaphysical *questions* and does not take us all the way to metaphysical speculation. This comes only when we mistake transcendental principles such as the Supreme Principle for 'constitutive' ones, that is, for true descriptions of the objects they refer to, whereas their only legitimate use is 'regulative,' that is, as hypothetical assumptions from which we derive research hypotheses. Put differently, we mistake 'subjective' principles (which are meant to direct our search for knowledge) for 'objective' principles (which truly describe reality). The tendency to make this mistake is what Kant calls 'transcendental illusion.' He explains it by appeal to a tacit assumption that underlies the speculative use of reason in metaphysics, namely, the assumption of 'transcendental realism.' As I will argue, 'transcendental realism' is here best understood as the claim that the structure of reality corresponds to that of rational thought, or, more generally, that the subjective conditions of thinking rationally about objects are conditions of the objects being thought about. Given this assumption, the regulative principles of reason that govern how we rationally *think* about objects must appear to be constitutive principles that characterize how those *objects* really are. I will maintain that Kant had good reason to think that an implicit commitment to transcendental realism is part of the 'nature' of 'universal human reason,' even though we may be more optimistic than Kant was as to whether we can rid ourselves of this implicit assumption.

0.4 Kant's Two Projects in the Transcendental Dialectic

It is an impression shared by many readers of the *Critique of Pure Reason* that this work falls into two major parts: a constructive one, comprising Kant's account of a priori cognition in the Transcendental Aesthetic and Analytic, and a destructive one, consisting in the demolition of traditional metaphysics in the Transcendental Dialectic. But such an impression can persist only if one ignores both the official structure of the book (which groups the Analytic and the Dialectic together as parts of Transcendental Logic) and the existence of the Doctrine of Method (officially the second main part of the book). Moreover, thinking of the Transcendental Dialectic as merely (or even predominantly) destructive obscures its important *constructive* strand, which is Kant's four-level Rational Sources Account. Kant's aim in the Transcendental Dialectic is not just to criticize traditional forms of metaphysics but equally to

show that they arise naturally out of indispensable and epistemically unproblematic employments of reason and are thus inscribed into the very structure of rational thinking itself. As Kant claims, it is his plan to "develop" the Transcendental Dialectic (and with it metaphysical speculation) "from its sources hidden deep in human reason" (A309/B366).

Thus, two projects, and two argumentative strands, are entwined in the Transcendental Dialectic: one constructive, the other destructive. Kant himself notes this at the end of the Transcendental Dialectic when he summarizes its results:

> The outcome of all dialectical attempts of pure reason not only confirms what we have already proved in the Transcendental Analytic, namely that all the inferences that would carry us out beyond the field of possible experience are deceptive and groundless, *but it also simultaneously teaches us this particular lesson*: that human reason has a natural propensity to overstep all these boundaries, and that *transcendental ideas are just as natural to it* as the categories are to the understanding, although with this difference, that just as the categories lead to truth … the ideas effect a mere, but irresistible, illusion, deception by which one can hardly resist even through the most acute criticism. (A642/B670; emphasis added)

The Transcendental Dialectic is thus both a critique of speculative metaphysics *and* an argument for the claim that there is a 'natural propensity' for metaphysical speculation that has its source in reason itself.[5] This latter 'lesson' is central to Kant's overall project for two reasons. First, Kant wants to criticize not just *some* historically prominent forms of metaphysics but *all possible* forms of *metaphysica specialis* (rational psychology, cosmology, and theology); for this, he needs to show that the proofs and inferences he criticizes are all there can possibly be, which presupposes that they are not contingent products of individual historical thinkers but realizations of a necessary rational structure. (I think that Kant's claim to have covered all possible forms of speculative metaphysics is questionable, and I will not try to defend it.)

But second, and more importantly, Kant aims to explain a fundamental feature of human existence – the urge to go, in thought, beyond the realm of empirical objects and to make claims about 'unconditioned' transcendent objects such as God, immortal souls, ultimate parts, and first causes. According to Kant, human beings have always had (and will always have) "a metaphysics of some kind" (*irgendeine Metaphysik*) (B21). They will always ask metaphysical questions, and certain answers to these questions will always appear plausible, or even irresistible, to them. Why is this the case? One central aim of the Transcendental Dialectic is to answer this question and thus to give a constructive, positive account of the metaphysical urge. If Kant is right, there

[5] These two sides of the Transcendental Dialectic have been clearly noted, e.g. in Klimmek 2005 and Pissis 2012.

is a sense in which metaphysics is rationally necessary (even though most of its pretensions may be unwarranted).[6]

In the long history of the reception of Kant's *magnum opus*, the first, destructive aspect of the Transcendental Dialectic has received much more attention than the second, constructive one, partly because it is more prominent in the text but also because its results must have struck Kant's early readers as much more challenging and important. From our present perspective, however, Kant's claim that we cannot prove the existence of God and the immortality of the soul appears less provocative than it did in Kant's own time; indeed, it has become part of enlightened common sense that, as Kant argues, one can only believe but not know that God exists. By contrast, Kant's claim that there is a natural tendency, grounded in reason itself, to ask metaphysical questions and a natural illusion that tempts us to believe that we can answer these questions remains provocative and exciting. I think it is time to pay more attention to this 'other side' of the Transcendental Dialectic, which consists in Kant's extended and highly complex argument for the Rational Sources Account. To the best of my knowledge, there has not yet been a book that focuses exclusively on this topic.[7]

This book appears at a time when the anti-metaphysical scruples that dominated Anglo-American philosophy in the second half of the twentieth century have largely subsided. In Kant scholarship, too, the pioneering work of Karl Ameriks (starting with Ameriks 1982) has led to a renewed interest in the

[6] Besides these two main projects in the Transcendental Dialectic, there are two further projects that are subordinated to, and dependent on, the main projects. The first is finding a positive use for the illusory concepts and principles that, according to the Rational Sources Account, arise from reason itself, which Kant attempts in the Appendix. The second consists in showing that Kant's critique of metaphysics does not undermine the logical and epistemic *possibility* of the immortality of the soul, freedom of the will, and God's existence, thus "mak[ing] room for *faith*" (Bxxx) and the postulates of pure reason.

[7] Michelle Grier's study on transcendental illusion (Grier 2001) covers some of the same territory but focuses on Kant's critique of metaphysics. Nikolai Klimmek reconstructs the "genesis of natural metaphysics" in the Transcendental Dialectic (Klimmek 2005: 2) but is primarily interested in Kant's system of transcendental ideas. Besides Grier's and Klimmek's books, there are a number of book-length studies and commentaries on the Transcendental Dialectic (e.g. Heimsoeth 1966–71; Bennett 1974; Pissis 2012), none of which discusses the rational sources of metaphysics for their own sake. R. Larnier Anderson's recent book also covers the Transcendental Dialectic but focuses on Kant's critique of rationalist metaphysics (Anderson 2015). Similarly, James Kreines, although he touches on the Rational Sources Account in his chapter on the Transcendental Dialectic, is primarily interested in how Kant argues for the limits of cognition (Kreines 2015: ch. 4). Susan Neiman's book on the unity of reason (Neiman 1994) overlaps with topics of the present work but does not isolate the Rational Sources Account as a single coherent strand in Kant's thinking (but see also Neiman 1995). In addition, there are important books on individual chapters and parts of the Transcendental Dialectic (e.g. Wood 1978; Ameriks 1982/2000; Malzkorn 1999; Rosefeldt 2000; and Falkenburg 2000, to name just a few) as well as studies with a more general focus (e.g. Guyer 1987; Gerhardt 2002; Höffe 2003; Allison 2004; Mohr 2004) that also cover some of the issues discussed in this book. While I have profited from all of these contributions, none of them focuses on Kant's Rational Sources Account in the way the present book does.

metaphysical aspects of Kant's critical thought (and an awareness of the continuities between the pre-critical and the critical Kant). This book fits into this general development in that it emphasizes the centrality of metaphysics to Kant's project of a critique of pure reason. As the final chapter of this book will show, however, I take Kant's critique of speculative metaphysics to be more radical than some proponents of the metaphysical reading of Kant will be happy to acknowledge, and this attitude will sometimes color my formulations in other parts of the book as well. Note, however, that my reconstruction of the Rational Sources Account in the main parts of this book is meant to be strictly neutral with respect to the extent to which speculative metaphysics is, or is not, compatible with Kant's critical philosophy, since its interest is not Kant's critique of metaphysics but his positive account of the sources of metaphysical thinking. I therefore hope that this book will be seen as a welcome addition to the literature by both proponents and critics of metaphysical interpretations of Kant.

0.5 Overview

This book will discuss the interrelations between two central aspects of Kant's philosophy, namely, his accounts of reason and of metaphysics. It will be obvious to anyone acquainted with Kant's work and the vast secondary literature surrounding it that such a project must be given strict limits. For one thing, I concentrate on Kant's views as developed in the *Critique of Pure Reason*, drawing on his other works only in order to clarify what Kant says in his *magnum opus*. This means that I will have to set aside the development of Kant's views over the course of his long philosophical career, including the pre-history of the Rational Sources Account in Kant's pre-critical writings and notes. I will also not be able to discuss the way in which Kant's understanding of reason and metaphysics develops after the B-edition of the first *Critique*. Moreover, even where the Transcendental Dialectic of the first *Critique* is concerned, I can engage with the chapters on the paralogisms, antinomies, and the arguments for God's existence, as well as the Appendix, only somewhat summarily, since a detailed treatment in each case would require a book of its own. Instead, I will concentrate on those aspects that are essential for an understanding of Kant's account of the rational sources of metaphysics.

The book has two main parts. Part I starts with an overview of Kant's accounts of reason and metaphysics and then offers a detailed interpretation of the first level at which Kant describes the slide from universal human reason to metaphysical speculation, namely, the transition from the Logical Maxim, which requires us to look for unconditioned cognitions, to the Supreme Principle, according to which if some conditioned object exists, there also exists something unconditioned. Chapter 1 offers a brief introduction to Kant's conceptions of reason and of metaphysics. Chapter 2 discusses the logical use

of reason, which is guided by the Logical Maxim and aims to transform our manifold cognitions about nature (both empirical and a priori) into a complete system of scientific knowledge and thus to achieve the 'unity of reason.' Chapter 3 turns to the Supreme Principle and the 'real' or 'transcendental' use of reason, which consists in tracking conditioning relations between objects (such as part–whole and substance–attribute relations) akin to what we today call 'metaphysical grounding' (see e.g. Correa and Schnieder 2012). I explain in detail what Kant, in the Supreme Principle, means by 'conditioned,' 'condition,' and the 'unconditioned' and argue that the relevant conditioning relations fall into three classes (corresponding to the three relational categories) without being species of a common genus. The main results of these chapters will be that three essential features of human reason (discursivity, iteration, and completeness) take us from ordinary employments of reason to a metaphysical search for the unconditioned.

Chapter 4 will then discuss the suggestive but cryptic passage (from the Introduction to the Transcendental Dialectic) in which Kant claims that the Logical Maxim 'cannot become' a principle of pure reason unless we 'assume' the Supreme Principle. Based on a close reading of the Appendix to the Transcendental Dialectic, I will argue that what is at issue here is a transition in two steps: first from the Logical Maxim to the regulatively used Supreme Principle, which is metaphysically harmless, and then from the latter to the constitutively used Supreme Principle, which carries with it a commitment to the existence of something unconditioned.

Finally, in Chapter 5, I explain why Kant thinks that both steps of this transition appear to be rationally necessary, even though only the first one really is. The first step *is* rationally necessary because we must go beyond the merely logical use of reason in order to approximate the 'unity of reason' (a complete system of scientific knowledge). The second step *appears* to be necessary because we tacitly assume that the structure of reality must correspond to the principles of reason. This assumption is what Kant calls 'transcendental realism,' and I explain how it follows from Kant's official definition of transcendental realism as the identification of appearances with things in themselves. Given transcendental realism, every regulative principle or concept of reason will appear to be constitutive of nature itself. This appearance is what Kant calls 'transcendental illusion.' Since we are rationally required to use the Supreme Principle regulatively (that is, as a heuristic hypothesis), given transcendental realism it seems to follow that it must necessarily be the case that, if there is something conditioned, there really *is* something unconditioned. As I will argue, Kant was right to assume that transcendental realism is a tacit background assumption that can plausibly be attributed to 'universal human reason.' In this way, Chapters 2–5 reconstruct the first, most basic level of Kant's Rational Sources Account.

At the same time, the transition from the Logical Maxim to the constitutive Supreme Principle provides a general template that is also at work on the other levels of Kant's argument: in moving from *logical* concepts and principles (which concern our cognitions) to real or *transcendental* ones (which concern objects), *transcendental realism* creates the illusion that the constitutive use of the latter is legitimate, since it conceals the difference between a legitimate *regulative* and an illegitimate *constitutive* use of concepts and principles of reason.

In Part II of the book, we will turn to the other three levels of Kant's Rational Sources Account to see how Kant applies this general template throughout the main text of the Transcendental Dialectic. In Chapter 6, we will discuss the second level, Kant's 'system of transcendental ideas' and the place of the 'metaphysical deduction' (subjective derivation) of those ideas. I will argue that the transcendental ideas are not derived from the forms of rational inferences (or the three possible relations a representation can have to its subject and object), as Kant seems to suggest. Rather, the transcendental ideas, which Kant calls 'inferred concepts,' are the result of 'necessary inferences of reason' – namely, the paralogisms, the cosmological arguments that lead up to the antinomies, and the one 'natural' argument for the existence of God that Kant discusses in Section Three of the Transcendental Ideal.

In Chapters 7 and 8, we turn to the third level of Kant's argument, which concerns the 'dialectical' (that is, illusory) inferences of reason and the derivation of the transcendental ideas. In Chapter 7, we discuss the paralogisms and the antinomies and how they allow us to derive specific transcendental ideas. I will argue that Kant provides us with a plausible account of why we tend to think of our souls as simple, persistent, and immaterial substances, and why contradictory claims about the world as a whole (e.g. that it is finite and that it is infinite; that it contains and does not contain simple parts and first causes) appear to be equally justified by rational arguments. As we will see, a reading of the paralogisms and antinomies from the perspective of Kant's Rational Sources Account reveals structural features of Kant's presentation that go unnoticed when the focus, as usual, is on Kant's critique of rational psychology and rational cosmology.

In Chapter 8, we then turn to rational theology and the derivation of the 'transcendental ideal' (the concept of the *ens realissimum*, or most real being) in Section Two of the Ideal of Reason chapter, which is widely considered to be obscure. However, if we read that section as part of Kant's Rational Sources Account (and as following the general pattern of a logical/transcendental transition plus a confounding of the regulative and constitutive uses of principles), many interpretative problems disappear. Concerning Kant's discussion of the three types of arguments for God's existence (ontological, cosmological, physicotheological) in Section Four of the Ideal, I will show that none of them features in Kant's argument for the Rational Sources Account. Instead, the

'natural' argument for God's existence is provided in Section Three of the Ideal of Reason. I will show that Kant makes a plausible case for the claim that the concept of an *ens realissimum* has its source in human reason and that a natural illusion can make us think that such a being must necessarily exist. Also in Chapter 8, we will address the fourth level of Kant's argument, discussed in the Appendix to the Transcendental Dialectic, and see why we tend to mistake regulative principles and ideas (which, according to Kant, are indispensable as heuristic devices in the scientific investigation of nature) for constitutive ones that appear to provide us with metaphysical insight into nature itself. The reason for this 'transcendental illusion,' as in all the other cases, is transcendental realism, which, however, can take a variety of forms.

Finally, in Chapter 9 we will round out our understanding of Kant's Rational Sources Account by asking how it relates to Kant's *critique* of speculative metaphysics. First, I will show that the different forms of transcendental realism appealed to in Kant's account are unified by a common core, which is the idea that reality must conform to the ways in which we necessarily represent it. Next, I argue that Kant's critique of speculative metaphysics is independent of any commitment to his own transcendental idealism. Rather, Kant's critique of the fallacies of rational psychology, cosmology, and theology in the Transcendental Dialectic requires only the rejection of transcendental realism, not the acceptance of transcendental idealism. Moreover, Kant's more general critique of *any* attempt to gain cognition of the unconditioned, or the supersensible, does not presuppose his transcendental idealism either and is instead based primarily on his claim that human cognition is limited to empirical objects. I also argue for a radical reading of Kant's account of transcendental ideas according to which they are cognitively defective ('without sense and significance') as long as we consider them only as part of metaphysical speculation and in abstraction from moral considerations. All in all, Kant mounts a compelling critique of the very kind of speculative metaphysics that his Rational Sources Account shows to be grounded in reason itself.

Each of the two parts of the book is followed by a Conclusion that contains an extensive summary of its main results and highlights the considerable plausibility, even from a present perspective, of Kant's account of the rational sources of metaphysics. The book closes with a brief Postscript on what I call Kant's 'practical metaphysics.' In the *Critique of Practical Reason* and elsewhere, Kant introduces three 'postulates of pure practical reason,' which concern God's existence, freedom of the will, and the immortality of the soul. I briefly discuss whether reason eventually finds in a practical mode what it had been looking for in a speculative mode, namely, the unconditioned. (The answer is: almost, but not quite.)