

# *Introduction to the Critique of Pure Reason*

---

PAUL GUYER AND ALLEN WOOD

Immanuel Kant's *Critique of Pure Reason* is one of the seminal and monumental works in the history of Western philosophy. Published in May 1781, when its author was already fifty-seven years old, and substantially revised for its second edition six years later, the book was both the culmination of three decades of its author's often very private work and the starting-point for nearly two more decades of his rapidly evolving but now very public philosophical thought. In the more than two centuries since the book was first published, it has been the constant object of scholarly interpretation and a continuous source of inspiration to inventive philosophers. To tell the whole story of the book's influence would be to write the history of philosophy since Kant, and that is beyond our intention here. After a summary of the *Critique*'s structure and argument, this introduction will sketch its genesis and evolution from Kant's earliest metaphysical treatise in 1755 to the publication of the first edition of the *Critique* in 1781 and its revision for the second edition of 1787.

## I.

### THE ARGUMENT OF THE CRITIQUE

**The strategy of the *Critique*.** In the conclusion to his second critique, the *Critique of Practical Reason* of 1788, Kant famously wrote, "Two things fill the mind with ever new and increasing admiration and awe the more often and more enduringly reflection is occupied with them: **the starry heavens above me and the moral law within me.**"- This motto could just as well have served for virtually all of Kant's philosophical works, and certainly for the *Critique of Pure Reason*. From the outset of his career, Kant had been concerned to resolve a number of the most fundamental scientific controversies of his epoch and to establish once and for all the basic principles of scientific knowledge of the world, thereby explaining our knowledge of the "starry heavens."

## Introduction

Almost as early in his career, Kant was intent on showing that human freedom, understood not only as the presupposition of morality but also as the ultimate value served and advanced by the moral law, is compatible with the truth of modern science. The *Critique of Pure Reason* was the work in which Kant attempted to lay the foundations both for the certainty of modern science and for the possibility of human freedom.

The book is complex, however, not just because of the complexity of Kant's own position, but also because he argues on several fronts against several different alternative positions represented in early modern philosophy generally and within the German Enlightenment in particular. In order to make room for his own dualistic defense of both modern science and human autonomy, Kant, like Descartes, Locke, and Hume, felt he had to rein in the pretensions of traditional metaphysics, which was represented for him by the school of Christian Wolff (1679–1754) and his followers, especially Alexander Gottlieb Baumgarten (1714–1762).<sup>2</sup> Their position, which Kant called "dogmatism," was compared in the Preface to the *Critique* to the despotic ministry of an absolute monarchy – Kant held dogmatism to be capricious, opinionated, faction-ridden and consequently unstable and open to the contempt of rational observers.

Yet Kant wanted to distinguish his own *critical* stance toward dogmatism from several other ways of rejecting it, which he regarded as themselves equally dangerous to the cause of reason. The first of these is *skepticism*, the position Kant took David Hume (1711–1776) to advocate.<sup>3</sup> Another position Kant rejected was *empiricism*,

the "way of ideas" described in John Locke's (1632–1704) *Essay concerning Human Understanding* (1790) as grounding knowledge solely on ideas acquired in the course of individual experience. Yet another philosophical stance Kant encountered was what he called *indifference* which did not reject metaphysical assertions themselves but did reject any attempt to argue for them systematically and rigorously. Here he had in mind a number of popular philosophers who were often in substantive agreement with dogmatists on metaphysical issues such as the existence of God and the immortality of the soul, but who were unconvinced by the scholastic subtlety of the dogmatists' propositions and proofs, holding instead that the beliefs on these matters that we need for the successful conduct of human life are simply given through "healthy understanding" or common sense.<sup>4</sup>

Yet while he attempted to criticize and limit the scope of traditional metaphysics, Kant also sought to defend against empiricists its underlying claim of the possibility of universal and necessary knowledge – what Kant called *a priori* knowledge, knowledge originating independently of experience, because no knowledge derived from any particular experience, or *a posteriori* knowledge, could justify a claim to universal and

## Introduction

necessary validity. He sought likewise to defend its scientific character against skeptics who dismiss its rigorous arguments as insufficient and against proponents of “common sense” who regard them as pedantic and superfluous. As Kant compared dogmatic metaphysicians to defenders of despotism, so he likened skeptics to nomads who abhor any form of permanent civil society and are prepared to disrupt or overthrow the monarchy of metaphysics, and Lockeans to calumniators who would foist a false and degrading genealogy on the monarch. Those who would pretend indifference to metaphysical inquiries he charged with being closet dogmatists, like supporters of a corrupt regime who scoff at its defects and feign ironic detachment from it but have no independent convictions of their own.

Kant’s position thus required him not only to undermine the arguments of traditional metaphysics but also to put in their place a scientific metaphysics of his own, which establishes what can be known *a priori* but also limits it to that which is required for ordinary experience and its extension into natural science. Kant therefore had to find a way to limit the pretensions of the dogmatists while still defending metaphysics as a science which is both possible (as was denied by the skeptics) and necessary (as was denied by the indifferentists). Thus Kant had to fight a war on several different fronts,<sup>5</sup> in which he had to establish the unanswerability of many metaphysical questions against both dogmatists and empiricists but also defend parts of the positions he was attacking, such as the possibility of *a priori* cognition of the fundamental principles of natural science, against both empiricists and skeptics. And while he wanted to prove to the indifferentists that a science of metaphysics is important, he also wanted to embrace part of their position, since he thought that in regard to some insoluble metaphysical questions, indeed the most important of them, we can defend a kind of commonsense belief – in God, freedom and immortality – because our moral outlook has an inescapable stake in them.

**The structure of the *Critique*.** This complex program led to the enormous complexity of the structure and argument of the *Critique of Pure Reason*. To many readers, the elaborate structure or “architectonic” of the *Critique* has been a barrier to understanding it, but a brief account of the origin of the main divisions of the book can illuminate its contents. Although these contents are profoundly original, Kant actually borrowed much of the book’s structure from well-known models. After the preface (which was completely rewritten for the second edition) and the introduction, the *Critique* is divided into two main parts, the “Doctrine of Elements” and the “Doctrine of Method.” This distinction is a variation on a distinction common in German logic textbooks between “general logic” and “special” or applied logic;<sup>6</sup> in Kant’s hands, it becomes a rubric to distinguish between his fundamental ex-

## Introduction

position of his theory of *a priori* cognition and its limits, in the “Doctrine of Elements,” and his own reflections on the methodological implications of that theory, under the rubric of the “Doctrine of Method,” where he provides contrasts between mathematical and philosophical proof and between theoretical and practical reasoning, as well as contrasts between his own critical method and dogmatic, empirical, and skeptical methods of philosophy.

The “Doctrine of Elements” in turn is divided into two main (although very disproportionately sized) parts, the “Transcendental Aesthetic” and the “Transcendental Logic,” the first of which considers the *a priori* contributions of the fundamental forms of our *sensibility*, namely space and time, to our knowledge, and the second of which considers the *a priori* contributions of the intellect, both genuine and spurious, to our knowledge. This division is derived from Baumgarten’s introduction of “aesthetics” as the title for the science of “lower” or “sensitive cognition” in contrast to logic as the science of higher or conceptual cognition;<sup>7</sup> at the time of writing the *Critique*, however, Kant rejected Baumgarten’s supposition that there could be a science of *taste* (what we now call “aesthetics”), and instead appropriated the term for his theory of the contribution of the forms of sensibility to knowledge in general.<sup>8</sup> After a brief explanation of the distinction between “general logic” and “transcendental logic” – the former being the basic science of the forms of thought regardless of its object and the latter being the science of the basic forms for the thought of objects (A50–57/B74–82) – Kant then splits the “Transcendental Logic” into two main divisions, the “Transcendental Analytic” and the “Transcendental Dialectic.” Kant uses this distinction, which derives from a sixteenth-century Aristotelian distinction between the logic of truth and the logic of probability, represented in eighteenth-century Germany by the Jena professor Joachim Georg Darjes (1714–1792),<sup>9</sup> to distinguish between the positive contributions of the understanding, working in cooperation with sensibility, to the conditions of the possibility of experience and knowledge (the “Transcendental Analytic”) and the spurious attempt of reason working independently of sensibility to provide metaphysical insight into things as they are in themselves (the “Transcendental Dialectic”). The “Transcendental Analytic” is in turn divided into two books, the “Analytic of Concepts” and the “Analytic of Principles,” the first of which argues for the universal and necessary validity of the pure concepts of the understanding, or the *categories*, such as the concepts of substance and causation, and the second of which argues for the validity of fundamental principles of empirical judgment employing those categories, such as the principles of the conservation of substance and the universality of causation.

The “Transcendental Dialectic” is also divided into two books, “On

## Introduction

the Concepts of Pure Reason” and “On the Dialectical Inferences of Pure Reason,” in which Kant explains how pure reason generates ideas of metaphysical entities such as the soul, the world as a whole, and God and then attempts to prove the reality of those ideas by extending patterns of inference which are valid within the limits of human sensibility beyond those limits. But it should be noted that the combination of the twofold division of the “Transcendental Analytic” into the “Analytic of Concepts” and “Analytic of Principles” with the main part of the Dialectic, the “Dialectical Inferences of Pure Reason,” replicates the traditional division of logic textbooks into three sections on *concepts*, *judgments*, and *inferences*.<sup>10</sup> Kant uses this structure to argue that the *concepts* of pure understanding, when applied to the forms of *sensibility*, give rise to sound principles of *judgment*, which constitute the heart of his critical metaphysics, but that *inferences* of pure reason performed without respect to the limits of sensibility give rise only to metaphysical illusion. The treatment of inferences is in turn divided into three sections, “The Paralogisms of Pure Reason,” “The Antinomy of Pure Reason,” and “The Ideal of Pure Reason,” which expose metaphysically fallacious arguments about the nature of the soul, about the size and origin of the world as a whole, and about the existence of God, respectively. These divisions are also derived from Kant’s predecessors: Wolff and Baumgarten divided metaphysics into “general metaphysics,” or “ontology,” and “special metaphysics,” in turn divided into “rational psychology,” “rational cosmology,” and “rational theology.” Kant replaces their “ontology” with the constructive doctrine of his own “Transcendental Analytic” (see A 247/B 303), and then presents his criticism of dogmatic metaphysics based on pure reason alone by demolishing the special metaphysics of rational psychology, cosmology, and theology.

Finally, Kant divides the “Doctrine of Method,” in which he reflects on the consequences of his demolition of traditional metaphysics and reconstruction of some parts of it, into four chapters, the “Discipline,” the “Canon,” the “Architectonic,” and the “History of Pure Reason.”<sup>11</sup> The first two of these sections are much more detailed than the last two. In the “Discipline of Pure Reason,” Kant provides an extended contrast between the nature of mathematical proof and philosophical argument, and offers important commentary on his own new critical or “transcendental” method. In the “Canon of Pure Reason,” he prepares the way for his subsequent moral philosophy by contrasting the method of theoretical philosophy to that of practical philosophy, and giving the first outline of the argument that runs through all three critiques, namely that *practical reason* can justify metaphysical *beliefs* about God and the freedom and immortality of the human soul although *theoretical reason* can never yield *knowledge* of such things. The last two parts of the “Doctrine of Method,” the “Architectonic of Pure Reason” and the

## Introduction

“History of Pure Reason,” recapitulate the contrasts between Kant’s own critical philosophical method and those of the dogmatists, empiricists, and skeptics with which he began, treating these contrasts in both systematic and historical terms. Indeed, although Kant himself never cared much about the history of philosophy as a scholarly discipline, in the few pages of his “History of Pure Reason” he outlined the history of modern philosophy as the transcendence of empiricism and rationalism by his own critical philosophy, the pattern that we still use, although of course we also have to add room to this pattern for the heirs and successors of Kant’s own philosophy.

With this analysis of the organization of the *Critique of Pure Reason* in hand, we now provide a brief resumé of its contents.

**“Introduction”: the idea of transcendental philosophy.** Although Kant himself often suggests that the negative side of his project, the critique of dogmatic metaphysics, is the most important, the *Critique* presents Kant’s positive doctrine of the *a priori* elements of human knowledge first. In the introduction, Kant argues that our mathematical, physical, and quotidian knowledge of nature requires certain judgments that are “synthetic” rather than “analytic,” that is, going beyond what can be known solely in virtue of the contents of the concepts involved in them and the application of the logical principles of identity and contradiction to these concepts, and yet also knowable *a priori*, that is, independently of any particular experience since no particular experience could ever be sufficient to establish the universal and necessary validity of these judgments. He entitles the question of how synthetic *a priori* judgments are possible the “general problem of pure reason” (B 19), and proposes an entirely new science in order to answer it (A 10–16/B 24–30).

This new science, which Kant calls “transcendental” (A 11/B 25), does not deal directly with objects of empirical cognition, but investigates the conditions of the possibility of our experience of them by examining the mental capacities that are required for us to have any cognition of objects at all. Kant agrees with Locke that we have no *innate* knowledge, that is, no knowledge of any particular propositions implanted in us by God or nature prior to the commencement of our individual experience.<sup>12</sup> But experience is the product both of external objects affecting our sensibility and of the operation of our cognitive faculties in response to this effect (A 1, B 1), and Kant’s claim is that we can have “pure” or *a priori* cognition of the contributions to experience made by the operation of these faculties themselves, rather than of the effect of external objects on us in experience. Kant divides our cognitive capacities into our receptivity to the effects of external objects acting on us and giving us sensations, through which these objects are given to us in empirical intuition, and our active faculty for relating the data of intuition by

## Introduction

thinking them under concepts, which is called understanding (A19/B33), and forming judgments about them. As already suggested, this division is the basis for Kant's division of the "Transcendental Doctrine of Elements" into the "Transcendental Aesthetic," which deals with sensibility and its pure form, and the "Transcendental Logic," which deals with the operations of the understanding and judgment as well as both the spurious and the legitimate activities of theoretical reason.

**"Transcendental Aesthetic": space, time, and transcendental idealism.** Despite its brevity – a mere thirty pages in the first edition and forty in the second – the "Transcendental Aesthetic" argues for a series of striking, paradoxical and even revolutionary theses that determine the course of the whole remainder of the *Critique* and that have been the subject of a very large proportion of the scholarly work devoted to the *Critique* in the last two centuries.<sup>13</sup> In this section, Kant attempts to distinguish the contribution to cognition made by our receptive faculty of sensibility from that made solely by the objects that affect us (A21–2/B36), and argues that space and time are pure forms of all intuition contributed by our own faculty of sensibility, and therefore forms of which we can have *a priori* knowledge. This is the basis for Kant's resolution of the debate about space and time that had raged between the Newtonians, who held space and time to be self-subsisting entities existing independently of the objects that occupy them, and the Leibnizians, who held space and time to be systems of relations, conceptual constructs based on non-relational properties inhering in the things we think of as spatiotemporally related.<sup>14</sup> Kant's alternative to both of these positions is that space and time are neither subsistent beings nor inherent in things as they are in themselves, but are rather only forms of our sensibility, hence conditions under which objects of experience can be given at all and the fundamental principle of their representation and individuation. Only in this way, Kant argues, can we adequately account for the necessary manifestation of space and time throughout all experience as single but infinite magnitudes – the feature of experience that Newton attempted to account for with his metaphysically incoherent notion of absolute space and time as the *sensorium dei* – and also explain the *a priori* yet synthetic character of the mathematical propositions expressing our cognition of the physical properties of quantities and shapes given in space and time – the epistemological certainty undercut by Leibniz's account of space and time as mere relations abstracted from antecedently existing objects (A22–5/B37–41, A30–2/B46–9).

Kant's thesis that space and time are pure forms of intuition leads him to the paradoxical conclusion that although space and time are *empirically real*, they are *transcendentally ideal*, and so are the objects given in them. Although the precise meaning of this claim remains subject to de-

## Introduction

bate,<sup>15</sup> in general terms it is the claim that it is only from the human standpoint that we can speak of space, time, and the spatiotemporality of the objects of experience, thus that we cognize these things not as they are in themselves but only as they appear under the conditions of our sensibility (A26–30/B42–5, A32–48/B49–73). This is Kant's famous doctrine of *transcendental idealism*, which is employed throughout the *Critique of Pure Reason* (and the two subsequent critiques) in a variety of ways, both positively, as in the “Transcendental Aesthetic” and “Discipline of Pure Reason,” to account for the possibility of synthetic *a priori* cognition in mathematics, and negatively, as in the “Transcendental Dialectic,” to limit the scope of our cognition to the appearances given to our sensibility, while denying that we can have any cognition of things as they are in themselves, that is, as transcendent realities constituted as they are independently of the constitution of our cognitive capacities.

**“Transcendental Analytic”: the metaphysical and transcendental deductions.** The longest and most varied part of the *Critique* is the “Transcendental Logic,” containing the two main divisions: the constructive “Transcendental Analytic,” which considers the *understanding* as the source of *a priori* concepts that yield *a priori* cognitions in conjunction with the forms of intuition already analyzed; and the primarily destructive “Transcendental Dialectic,” which investigates the faculty of *reason*, in the first instance as a source of illusory arguments and metaphysical pseudo-sciences, although in the end also as the source of valuable regulative principles for the conduct of human inquiry and practical reasoning. The “Transcendental Analytic,” as we saw, is in turn divided into two books, the “Analytic of Concepts,” dealing with the *concepts* of the understanding, and the “Analytic of Principles,” concerning the *principles* of the understanding that arise from the application of those concepts to the forms of intuition.

In the “Analytic of Concepts,” Kant presents the understanding as the source of certain concepts that are *a priori* and are conditions of the possibility of any experience whatever. These twelve basic concepts, which Kant calls the *categories*, are *fundamental concepts of an object in general*, or the forms for any particular concepts of objects, and in conjunction with the *a priori* forms of intuition are the basis of all synthetic *a priori* cognition. In an initial section of the “Transcendental Analytic” (A66–81/B91–116), which he named in the second edition of the *Critique* the “metaphysical deduction” of the categories (B159), Kant derives the twelve categories from a table of the twelve *logical functions* or forms of judgments, the logically significant aspects of all judgments. Kant's idea is that just as there are certain essential features of all judgments, so there must be certain corresponding ways in which we form the concepts of objects so that judgments may be about objects.

## Introduction

There are four main logical features of judgments: their *quantity*, or the scope of their subject-terms; the *quality* of their predicate-terms, whose contents are realities and negations; their *relation*, or whether they assert a relation just between a subject and predicate or between two or more subject-predicate judgments; and their *modality*, or whether they assert a possible, actual, or necessary truth. Under each of these four headings there are supposed to be three different options: a judgment may be universal, particular or singular; affirmative, negative or infinite; categorical, hypothetical or disjunctive; and problematic, assertoric, or apodictic. Corresponding to these twelve logical possibilities, Kant holds there to be twelve fundamental categories for conceiving of the quantity, quality, relation, and modality of objects (A70/B95, A80/B106). The plausibility of Kant's claim that there are exactly twelve logical functions of judgment and twelve corresponding categories for conceiving of objects has remained controversial since Kant first made it.<sup>16</sup>

Even if Kant establishes by this argument that we have certain concepts *a priori*, it is a more ambitious claim that all of these concepts apply universally and necessarily to the objects that are given in our experience. Kant takes on this more ambitious project in the "Transcendental Deduction of the Categories," the chapter which he says in the first edition of the *Critique* cost him the most labor (Axvi), but which he then rewrote almost in its entirety for the second edition (A84–130/B116–69) after other attempts in the intervening works, the *Prolegomena to Any Future Metaphysics* (1783) and *Metaphysical Foundations of Natural Science* (1786). In both versions of the *Critique*, although not in the intervening works, Kant centers his argument on the premise that our experience can be ascribed to a single identical subject, via what he calls the "transcendental unity of apperception," only if the elements of experience given in intuition are synthetically combined so as to present us with objects that are thought through the categories. The categories are held to apply to objects, therefore, not because these objects make the categories possible, but rather because the categories themselves constitute necessary conditions for the representation of all possible objects of experience. Precisely what is entailed by the idea of the unity of apperception, however, and what the exact relation between apperception and the representation of objects is, are obscure and controversial, and continue to generate lively philosophical discussion even after two centuries of interpretation.<sup>17</sup>

**Principles of pure understanding.** Even if the transcendental deduction does establish that the categories do apply to all possible data for experience, or (in Kant's terms) all manifolds of intuition, it does so only abstractly and collectively – that is, it does not specify how each category applies necessarily to the objects given in experience or show

## Introduction

that all of the categories must be applied to those objects. This is Kant's task in Book II of the "Transcendental Analytic," the "Analytic of Principles." This book is in turn divided into three chapters, "The Schematism of the Pure Concepts of the Understanding," the "System of All Principles of Pure Understanding," and "On the Ground of the Distinction of All Objects in General into Phenomena and Noumena." In the first of these chapters Kant shows how the logical content of the categories derived from the metaphysical deduction is to be transformed into a content applicable to the data of our senses; in the second, he demonstrates principles of judgment showing that all of the categories must be applied to our experience by means of arguments that are sometimes held to prove the objective validity of the categories independently of the prior transcendental deduction; and in the third chapter Kant draws out the consequences of the preceding two, arguing that because the categories have a determinate use only when applied to spatiotemporal data and yet the forms of space and time themselves are transcendently ideal, the categories also have a determinate cognitive use only when applied to appearances ("phenomena"), and therefore that by means of the categories things as they are in themselves ("noumena") might be *thought* but not *known*.

In the "Schematism," Kant argues that the categories, whose content has thus far been derived solely from the logical structure of judgments, must be made applicable to objects whose form has thus far been specified solely by the pure forms of space and time. He argues that this can be done by associating each category with a "transcendental schema," a form or relation in intuition that is an appropriate representation of a logical form or relation. In particular, Kant argues that each category must be associated with a *temporal* schema, since time is the form of every sensible intuition whatever, while space is the form of outer intuitions only. For example, the schema of the logical conception of ground and consequence is the concept of *causality* as rule-governed temporal *succession*: the concept of a cause, as opposed to that of a mere ground, is the concept of "the real upon which, whenever it is posited, something else always follows," or "the succession of the manifold insofar as it is subject to a rule" (A 144/B 183). As Kant will make clearer in the second edition, however, the subsequent chapter on the "Principles" will show that although the content of the transcendental schemata for the categories may be explicated in purely temporal terms, the *use* of these schemata in turn depends upon judgments about the *spatial* properties and relations of at least some objects of empirical judgment. Thus the argument of the "Analytic of Principles" as a whole is that the categories both must and can only be used to yield knowledge of objects in space and time. The principles expressing the universal and necessary application of the categories to objects given in space and time are precisely

## Introduction

~~the synthetic *a priori* judgments that are to be demonstrated by Kant's critical replacement for traditional metaphysics.~~

In the second chapter of the "Analytic of Principles," the "System of All Principles," Kant organizes the principles of pure understanding under four headings corresponding to the four groups of categories. For each of the first two groups of categories, those listed under "Quantity" and "Quality," Kant supplies a single "mathematical" principle meant to guarantee the application to empirical objects of certain parts of mathematics, which are in turn supposed to be associated with certain parts of the logic of judgment. The first principle, under the title "Axioms of Intuition," guarantees that the *a priori* mathematics of ~~extensive magnitudes, where wholes are measured by their discrete parts,~~ applies to empirical objects because these are given in space and time which are themselves extensive magnitudes (A162–6/B202–7). The general implication of this argument is that the empirical use of the logical quantifiers (one, some, all) depends on the division of the empirical manifold into distinct spatiotemporal regions. The second principle, under the title of the "Anticipations of Perception," guarantees that the mathematics of *intensive magnitudes* applies to the "real in space," or that properties such as color or heat, or material forces such as weight or impenetrability, must exist in a continuum of degrees because our sensations of them are continuously variable (A166–76/B207–18). Here Kant's argument is that since the use of the logical functions of affirmation and negation is dependent on the presence or absence of sensations that come in continuously varying degrees, the empirical use of the categories of "Quality" is connected with the mathematics of intensive magnitudes in a way that could not have been predicted from an analysis of the logical content of these categories themselves (another example of how a synthetic *a priori* rather than merely analytic judgment arises).

Switching from "mathematical" to "dynamical" principles, the third section of the "System," the "Analogies of Experience," concerns the necessary relations among what is given in space and time, and thus gives expression to the necessary conditions for the application of the categories of "Relation" to empirical objects. Many interpreters consider this the most important section of the *Critique*. In the first analogy, Kant argues that the unity of time implies that all change must consist in the alteration of states in an underlying substance, whose existence and quantity must be unchangeable or conserved (A182–6/B224–32). In the second analogy, Kant argues that we can make determinate judgments about the objective succession of events as contrasted to merely subjective successions of representations only if every objective alteration follows a necessary rule of succession, or a causal law (A186–211/B232–56). In the third analogy, Kant argues that determinate judgments

## Introduction

that objects (or states of substance) in different regions of space exist simultaneously are possible only if such objects stand in the mutual causal relation of community or reciprocal interaction (A211–15/B256–62). The second analogy is generally supposed to supply Kant's answer to Hume's skeptical doubts about causality, while the third analogy is the basis for Kant's refutation of Leibniz's rejection of real interaction between independent substances—an essential thesis of Leibniz's "monadology." In particular, both what the second analogy is intended to prove and how the proof is supposed to proceed have been matters of exegetical controversy; they have been disputed almost as intensely as the philosophical question whether Kant's reply to Hume is successful.

In the first edition of the *Critique*, the final section of the "System of Principles," the "Postulates of Empirical Thought," provides conditions for the empirical use of the modal categories of possibility, existence, and necessity, and argues that our determinate use of the categories of both possibility and necessity is in fact confined to the sphere of the *actual*, that is, that which is actually given in experience (A218–35/B265–74, 279–87). In the second edition, however, Kant inserted a new argument, the "Refutation of Idealism" (B274–9), which attempts to show that the very possibility of our consciousness of ourselves presupposes the existence of an external world of objects that are not only represented as spatially outside us but are also conceived to exist independently of our subjective representations of them. Although the implications of this argument have been intensely debated, it seems to confirm Kant's claim in the *Prolegomena to Any Future Metaphysics* that his "transcendental idealism" is a "critical" or "formal" idealism that, unlike traditional idealism, implies the subjectivity of space and time as forms of intuition without denying the real *existence* of the objects distinct from ourselves that are represented as being in space and time.<sup>18</sup>

In the third chapter of the "Analytic of Principles," on phenomena and noumena, Kant emphasizes that because the categories must always be applied to data provided by sensibility in order to provide cognition, and because the data of sensibility are structured by the transcendently ideal forms of intuition, the categories give us knowledge only of things as they appear with sensibility ("phenomena," literally "that which appears"). Although through pure understanding (*noûs* in Greek) we may *think* of objects independently of their being given in sensibility, we can never *cognize* them as such non-sensible entities ("noumena," literally "that which is thought") (A235–60/B294–315). The meaning of Kant's use of the term "phenomena" is self-evident, but the meaning of "noumena" is not, since it literally means not "things as they are independently of appearing to us" but something more like "things as they are understood by pure thought." Yet Kant appears to deny that the human understanding can comprehend things in the latter way. For

## Introduction

this reason, Kant says it is legitimate for us to speak of noumena only “in a negative sense,” meaning things as they may be in themselves independently of our representation of them, but not noumena “in a positive sense,” which would be things known through pure reason alone. A fundamental point of the *Critique* is to deny that we ever have knowledge of things through pure reason alone, but only by applying the categories to pure or empirical data structured by the forms of intuition.

At this point in the *Critique* Kant has completed the largest part of his constructive project, showing how synthetic *a priori* principles of theoretical cognition are the necessary conditions of the application of the categories to sensible data structured by the pure forms of intuition. The next part of his argument is the critical demonstration that traditional metaphysics consists largely of illusions arising from the attempt to acquire knowledge of all things (the soul, the world as a whole, and God) as they are in themselves by the use of reason alone regardless of the limits of sensibility. The bulk of this argument is reserved for the “Transcendental Dialectic,” but Kant makes a start on it with the interesting appendix that completes the “Transcendental Analytic” entitled the “Amphiboly of Concepts of Reflection” (A160–92 / B316–49). In this appendix Kant presents his criticism of Leibniz’s monadology by arguing that through a confusion (or “amphiboly”) Leibniz has taken mere features of concepts through which we think things, specifically concepts of comparison or reflection such as “same” and “different” or “inner” and “outer,” which are in fact never applied directly to things but only applied to them through more determinate concepts, as if they were features of the objects themselves. Kant thereby rejects the Leibnizian-Wolffian account of such metaphysical concepts as essence, identity, and possibility, and reinforces his own insistence that empirical individual judgments of real possibility require sensible conditions in addition to logical intelligibility and non-contradictoriness.

**The “Transcendental Dialectic”: the critique of metaphysics.** The second division of the “Transcendental Logic” turns to the main destructive task of the *Critique of Pure Reason*, and that which gives it its name, the task of discrediting dogmatism and displaying the limits of metaphysics. The “Transcendental Analytic” has prepared the way for this critique of traditional metaphysics and its foundations by its argument that synthetic *a priori* principles can be established only within the limited domain of sensible experience. But Kant’s aim in the “Dialectic” is not only to show the failure of a metaphysics that transcends the boundaries of possible experience. At the same time, he also wants to demonstrate that the questions that preoccupy metaphysics are inevitable, and that the arguments of metaphysics, although deceptive, should not be dismissed without sympathetic comprehension (as they are by the traditional skeptic). Kant argues that they tempt us for gen-

## Introduction

nine reasons, inherent in the nature of human reason itself, and when these grounds are properly understood they can be put to good use for the causes of both human knowledge and human morality. This argument is the basis for Kant's theory of the regulative use of the ideas of reason in scientific inquiry, which Kant first suggests in the final appendix to the "Transcendental Dialectic" and then elaborates in the *Critique of Judgment*, and for his theory of the foundation of morality in the practical use of pure reason, which he first describes in the "Doctrine of Method" and elaborates in many subsequent works, but especially in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*.

The Leibnizian-Wolffian tradition, as presented in Alexander Gottlieb Baumgarten's *Metaphysica* (first edition, 1738), which Kant used as the textbook for his lectures on metaphysics for virtually his entire career, was divided into four parts: ontology, psychology, cosmology, and theology. The "Transcendental Aesthetic" and "Analytic" are Kant's critical replacement for traditional ontology. The "Transcendental Dialectic," however, is dedicated to arguing that the other three parts of the rationalist system are pseudo-sciences founded on inevitable illusions of human reason attempting to extend itself beyond the limits of sensibility. Kant does not present the three rationalistic pseudo-sciences as mere historical artifacts, but attempts to display them as inevitable products of human reason by associating them with the unconditioned use of the three traditional forms of syllogism: categorical, hypothetical, and disjunctive. Seeking the unconditioned subject to which all our thoughts relate as predicates, we generate the idea of the soul as a simple, non-empirical substance; seeking the unconditioned in respect of any of several hypothetical series arising in the world (of composition or extension, of decomposition or division, of cause and effect) leads to ideas such as that of a first event in time, an outer limit to space, a simple substance and a first cause. Finally, Kant derives the idea of a most real being or God as the ideal ground of the real properties constituting all other things. Kant's overall argument is that although these rationalist doctrines are inevitable illusions they are still pseudo-sciences, and must give way to doctrines remaining within the limits of sensibility: rational psychology gives way to empirical psychology, which Kant expounded in his lectures in the form of "anthropology"; rational cosmology gives way to the metaphysical foundations of natural science, which Kant derives by adding the sole empirical concept of motion to the principles of judgment; and rational theology gives way to what Kant will call moral theology, the doctrine that God and immortality are postulated, along with freedom of the will, solely as conditions of the possibility of human morality.

The opening book of the "Transcendental Dialectic" is therefore a

## Introduction

derivation and even a limited defense of the *transcendental ideas*, such as the immortal soul, free will, and God, with which dogmatic metaphysics has always been preoccupied (A293–338/B349–96). *Reason*, traditionally thought to be the highest of our cognitive faculties, has a “logical use” in which it simply draws inferences from principles, but also a “real use” in which it seeks to base series of ordinary inferences, such as those from cause to effect, in ultimate, foundational principles, such as the idea of an uncaused first cause. The ideas of such ultimate principles are generated *a priori* by the faculty of reason when it seeks, through regressive syllogistic reasoning, for what is *unconditioned* in respect of the objects given in experience, according to the principles of understanding that govern these objects. In particular, it is the three categories of *relation* when used without regard to the limits of sensibility that give rise to the chief ideas of metaphysics: the concept of substance giving rise to the idea of the soul as the ultimate subject, the concept of causation giving rise to the idea of the world-whole as a completed series of conditions, and the concept of community giving rise to the idea of God as the common ground of all possibilities. Kant suggests that each of the three relational categories gives rise to a distinctive form of syllogistic inference, series of which can only be terminated by the idea of an unconditioned ground, but also that the attempt to acquire knowledge by means of the relational categories without sensibility gives rise directly to the idea of an unconditioned subject, series, and set of all possibilities.

The second and by far the larger book of the “Dialectic” expounds “The Dialectical Inferences of Pure Reason” in great detail. The errors of rational psychology are diagnosed under the rubric of “The Paralogisms of Pure Reason,” those of rational cosmology under the rubric of “The Antinomy of Pure Reason,” and those of rational theology under the rubric of “The Ideal of Pure Reason.”

~~The “Paralogisms.” Rational psychology is the topic of the “Paralogisms” (or fallacious inferences) of pure reason, which argue invalidly from the formal unity, simplicity, and identity of the *thought* of the subject of thinking or the “I” to the conclusion that the *soul* is a real and simple (hence indestructible) substance that is self-identical throughout all experience (A341–66). In the first edition, the “Paralogisms” included a fourth part, which defends the reality of external appearance in space simply by reducing objects in space to one form of immediate representation (A366–405). This response to idealism appears to provide only a Pyrrhic victory over it, which provoked charges of Berkeleyanism against Kant, and was therefore replaced in the second edition with the “Refutation of Idealism,” which as we saw argues for the real existence of objects in space and time although for the transcendental ideality of their spatial and temporal form. In the second~~

## Introduction

~~edition, the entire chapter on the paralogisms was rewritten and simplified (B 406–22); to fill the place of the superseded fourth paralogism, Kant adds an argument that his dualism of appearance and reality undercuts the traditional dualism of mind and body, with its problem about the possibility of interaction between two fundamentally distinct kind of substances, by opening up the possibility that both mind and body are different appearances of some single though unknown kind of substance.~~

~~The “Antinomies.” The longest and most painstaking part of the “Transcendental Dialectic” is the “Antinomy of Pure Reason,” which deals with the topics of rational cosmology (A405–583/B432–611), indeed, as we will show below, Kant originally thought that all of the errors of metaphysics could be diagnosed in the form of these antinomies. Here Kant argues that reason’s natural illusions are not merely revealed by subtle philosophical analysis but unavoidably manifest themselves in the form of actual contradictions each side of which seems naturally plausible. Kant argues that unless we accept the transcendental idealist distinction between appearances and things in themselves, we will be committed to accepting mutually incompatible arguments, arguments both that there must be a first beginning of the world in time and that there cannot be, that there must be limits to the world in space and that there cannot be (the two halves of the first antinomy), both that there must be a simple substance and that there cannot be (the second antinomy), both that there must be at least one first or uncaused cause and that there cannot be (the third antinomy), and that there must be a being whose necessary existence is the ground of all contingent beings and that there can be no necessary being (the fourth antinomy).~~

~~The only way of resolving these contradictions, Kant argues, is by accepting that the natural world is a realm of appearances constituted by the application of the categories to sensible intuitions, and not a realm of things in themselves. Regarding the first two antinomies, which he calls “mathematical” antinomies because they have to do with size and duration, Kant argues that there is no fact of the matter about the size of the world as a whole, because the natural world is never present in experience as a whole, but rather is given to us only through the progressive or regressive synthesis of spatiotemporal intuitions. We can always proceed *indefinitely* far in the progressive composition of spaces and times into ever larger or longer realms or in the regressive decomposition of space and time into ever smaller regions, but we can never reach a beginning or an end to such series, as would be possible if they were finite, nor complete any synthesis of them as infinite either. Both sides of the mathematical antinomies, therefore, turn out to be false, because both rest on the common—and false—assumption that the world is given independently of our ongoing synthesis in its representation,~~

## Introduction

and that it therefore has a determinate magnitude, which must be either finite or infinite. For the third and fourth antinomies, which he calls “dynamical” because they have to do with the causation of the world and its events, Kant proposes a different solution. Here he argues that both sides may be true, if the denial of a free cause or necessary being is restricted to the natural and sensible world and their affirmation is taken to refer to what might exist in a noumenal or supersensible world of things in themselves. Just as his thinking about the antinomies generally shaped his thinking about the structure and outcome of the entire “Transcendental Dialectic,” so Kant’s resolution of the third antinomy will go on to play an important role in his moral philosophy and in his ultimate account of the relation between theoretical and moral philosophy.

**The “Ideal of Pure Reason.”** Rational theology, the third and last of the metaphysical pseudo-sciences, is taken up by Kant in the final chapter of the “Transcendental Dialectic” (A 567–642/B 595–670). If an “idea” is a pure concept generated by reason, then an “ideal” is the concept of an *individual thing* as exemplifying an idea of pure reason. It would not be natural to think of the idea of the soul, for example, as giving rise to an ideal, because we naturally think there are many souls; but it is natural (at least in the Judaeo-Christian tradition) to think of the idea of God as the idea of a single thing, and thus the idea of God is the *ideal* of pure reason. Kant argues for the inevitability of the idea of God as an *ens realissimum*, or supreme individual thing possessing all realities or perfections and thus also grounding all the possibilities realized by other particular things. Much of Kant’s argument here makes use of a line of thought he developed nearly twenty years before the publication of the *Critique in The Only Possible Ground of Proof for a Demonstration of the Existence of God* (1763). But now Kant subjects to withering criticism his own earlier attempt to prove the existence of God as such an *ens realissimum* as well as the other traditional attempts to prove the existence of God, which were already criticized in Kant’s earliest philosophical writing, the *New Elucidation of the First Principles of Metaphysical Cognition* (1755) as well as in *The Only Possible Ground*.

Kant organizes the traditional proofs of the existence of God (without attempting to explain why there should only be these three) into the *ontological* proof, based solely on the *concept* of God, the *cosmological* proof, based on the sheer fact of the *existence* of a world, and the *physico theological* proof, based on the particular constitution of the actual world, especially its alleged exhibition of purposive design. The first of these is Kant’s representation of the proof favored by St. Anselm and revived by Descartes; the second is his name for an argument from contingent existents to their necessary ground favored by Wolff and his followers; and the third is what Kant calls the argument

## Introduction

from design favored by so many thinkers of the early Enlightenment, especially in Britain (where Hume had already subjected it to trenchant criticism in his *Dialogues concerning Natural Religion*, which, because of the delay of their translation into German, Kant had not yet seen at the time he published the *Critique*). First Kant attacks the ontological argument, holding that since existence is not a property and therefore not itself a perfection, it cannot be included among the contents of the idea of God, and cannot be inferred from that idea alone. Instead, Kant argues, the existence of an object is always the presupposition of the truth of any assertion about it, and cannot itself be assumed for the proof of such an assertion. Kant then argues that even if the cosmological and physico-theological proofs could establish the existence of some necessary and purposive being (which they cannot), they still could not establish the existence of a supremely perfect Deity unless the ontological proof also succeeded. Since the ontological proof is unsound, the entire metaphysical enterprise of proving the existence of God – as an object of theoretical cognition – must be given up as hopeless.

**Regulative use of the ideas.** The outcome of the “Transcendental Dialectic,” therefore, seems to be entirely negative. This is a misleading conclusion, however. In an appendix to the “Dialectic,” Kant begins a limited rehabilitation of the ideas of traditional metaphysics by arguing that the ideas of reason have an important function in the conduct of natural science if they are understood *regulatively*, that is, if they are taken to represent not metaphysical beings or entities whose reality is supposed to be demonstrable, but rather goals and directions of inquiry that mark out the ways in which our knowledge is to be sought for and organized. This is true of the idea of a simple soul, which stimulates us to search for a unified psychology; of the idea of a complete world-whole, which leads us constantly to expand the domain of our scientific investigations; and above all of the idea of God, for regarding the world as if it were the product of a highest intelligence leads us to look for the maximum in order and connectedness, which is beneficial for the organization of whatever empirical knowledge we do acquire. This argument, which Kant continues in the *Critique of Judgment*, is the first of Kant’s constructive arguments that reason can be misleading but if wisely used is far from idle or even unnecessary. Kant’s second constructive argument about reason, that its ideas have a profound practical use for the guidance and regulation of conduct, is begun in the final part of the *Critique*, the “Doctrine of Method.”

**“The Doctrine of Method.”** The second major division of the *Critique*, the “Doctrine of Method,” tends to be neglected by its readers, perhaps because the “Doctrine of Elements” is so long and the arguments already surveyed are so exhaustive. But the “Doctrine of

## Introduction

Method," in which Kant reflects upon the potential and the limits of his critical philosophy by comparing it with other methods—he compares the method of philosophy with the method of mathematics, the method of theoretical philosophy with the method of practical philosophy, and the method of critical philosophy with the methods of dogmatic, empirical, and skeptical philosophy—includes some extremely important discussions. Its first chapter, the "Discipline of Pure Reason," provides Kant's most mature treatment of the difference between philosophy and mathematics, arguing that both provide synthetic *a priori* cognition, but that mathematics provides determinate answers to its problems because its objects can be constructed in pure intuition, whereas philosophy provides only general principles because what it can construct are the conditions of possibility for the experience of objects, not particular objects (A712–38/B740–69). Then it provides an ardent defense of freedom of public communication as well as of open mindedness in the discussion of metaphysical issues, arguing that the very existence of reason itself depends on the free give and take of controversy between rational beings, which requires the liberty to come to one's own conclusions honestly and to express them openly to others (A738–69/B766–97). This discussion presages Kant's impassioned defense of freedom of thought in his political writings of the 1790s. The chapter concludes with a discussion of the contrasting roles of hypotheses in science and philosophy (A769–82/B798–810) and then with a reflection upon his own style of philosophical argumentation, what he calls "transcendental proofs" (A782–94/B810–22).

The second chapter of the "Doctrine of Method," the "Canon of Pure Reason," contrasts the epistemological status of theoretical cognition with that of the principles and presuppositions of practical reason, or morality, and in so doing provides Kant's most systematic discussion of moral philosophy prior to the *Groundwork of the Metaphysics of Morals* (1785) and Kant's first systematic statement of his argument for rational faith in God on moral grounds (A795–831/B823–59), an argument that Kant was to restate and refine in the subsequent two critiques and to continue to work on until the end of his life. The third chapter, the "Architectonic of Pure Reason," continues the discussion of the contrast between philosophy and other forms of cognition, such as historical knowledge, as well as of the contrast within philosophy between theoretical and practical reason (A832–51/B860–79), while the final chapter of the "Doctrine of Method," and of the whole *Critique*, the "History of Pure Reason," orients the critical philosophy clearly in relation to the competing positions of dogmatism, empiricism, skepticism, and indifferentism, the discussion of which had opened the *Critique* (A852–56/B880–84). For all its brevity, this section has had considerable influence on subsequent conceptions of the history of philosophy.

## Introduction

### II.

#### THE MESSAGE OF THE CRITIQUE

The *Critique of Pure Reason* is complex and many-sided. Both its overall message and its meaning for the subsequent history of philosophy defy any easy summary. The *Critique* has perhaps most often been seen as marking out a third way that combines the virtues, while avoiding the pitfalls, of both the “rationalism” of Descartes and Leibniz and the “empiricism” of Locke and Hume. This way of reading the *Critique*, however, even though to some extent suggested by Kant himself, depends on a simplified reading of the history of modern philosophy and at the very least on an incomplete assessment of the strengths and weaknesses of Kant’s modern predecessors. Less controversial is the observation that the *Critique*’s main intention is to find a middle way between traditional metaphysics, especially its attempts to bolster a theistic view of the world with *a priori* rational arguments, and a skepticism that would undercut the claims of modern natural science along with those of religious metaphysics.

We see this clearly in the way that Kant defines the position of critical philosophy in contrast to dogmatism, empiricism, skepticism, and indifferentism. He seeks to carve out for theoretical philosophy a significant but limited domain, distinct from that of empirical knowledge and the opinions of common sense, but excluding the exaggerated claims that have brought metaphysics into disrepute. In this way, the *Critique of Pure Reason* belongs to a main tradition in modern philosophy, beginning with Descartes, that tries to provide an *a priori* philosophical foundation for the methods and broad features of a modern scientific view of nature by an examination of the suitability of human cognitive faculties for the kind of knowledge of nature that modern science aims to achieve. At the same time, Kant tries to save precisely what the dogmatic metaphysicians cannot, by connecting the claims of religious metaphysics not to the sphere of theory but to the sphere of moral practice, and, in the famous words of the second-edition preface, by limiting knowledge in order to make room for faith (Bxxx). But Kant tries to accomplish all these goals, especially the last, in an authentically *Enlightenment* manner, always giving first place to our rational capacity to reflect on our cognitive abilities and achievements, to correct them, and to subject the pretensions of reason to *self-limitation*, so that human reason itself retains ultimate authority over all matters of human knowledge, belief, and action. The ultimate autonomy of human thought lies in the fact that it neither can nor must answer to any authority outside itself.

The originality of the *Critique* can be indicated by focusing on the way it attempts simultaneously to resolve two of the most intractable

## Introduction

problems of early modern philosophy, the simultaneous vindication of the principle of universal causality and of the freedom of the human will. The great idea of the *Critique of Pure Reason* is that the very thing that explains the possibility of our knowledge of the fundamental principles grounding a scientific view of nature is also the key to the possibility of our freedom in both intention and action, which seems threatened by the rule of causality in that natural world. Kant argues that the principles of the scientific worldview can be known with certainty because they express the structure of our own thought. They are therefore conditions of the possibility of our experience, which we impose upon the raw data of sensation. Thus, there is a sense in which certitude about the principles of science is possible only because of human autonomy: we are not merely passive perceivers of sensible information flowing into us from external objects, but also cognitive agents who structure what we perceive in accordance with the necessary conditions of our active thought. Thus Kant argues that we can be certain of the fundamental principles of science – above all the universal law of causation, the assumption underlying all scientific inquiry that every event has a cause and can therefore be explained in accordance with a law of nature – precisely because this law is a condition of the possibility of the thought that we must impose upon our perceptions in order to have any experience at all.

In the seventeenth and eighteenth centuries, the principle of causation had been put into ever more successful use by practicing scientists, but at the same time doubt had been cast upon it by philosophers. First the principle had been supported upon theological foundations by Descartes and his follower Nicolas Malebranche, and then reduced to a mere phenomenon, as by Leibniz, or finally exposed by Hume as simply the result of mere custom. Kant, however, argues that a genuine necessary connection between events is required for their objective succession in time, and that the concept of causality in which this connection is expressed is imposed on experience by our own thought as an indispensable condition of its possibility. The human understanding, therefore, is the true lawgiver of nature, and the successes of modern science are due to its conduct of its inquiries in accordance with a plan whose ground lies *a priori* in the structure of human thought (Bxii–xviii). At the same time, nature is to be regarded as essentially an object of human sensation and thinking, and the validity of the causal principle is to be restricted to the world as it appears under the conditions of our experience of it. In this way, the same account that guarantees the certitude of the principle of causation also guarantees the freedom of the human will, which is precisely what was typically thought to be excluded by the universality of causation.

According to Kant, if we understand the principle of causality and the

## Introduction

fundamental principles of the scientific worldview as products of our own thought imposed upon experience, this leaves open the possibility of a radical self-determination of human action when the human will is considered not as it appears but as it is in itself. In later works, such as the *Critique of Practical Reason* (1788) and the *Religion within the Boundaries of Mere Reason* (1793), Kant completes this theory with the further argument that only the inexorable awareness of our obligation to live up to the moral law, which is given spontaneously by our own reason and which we all acknowledge (even if only in the breach), can prove the reality of our freedom, which is the necessary condition of the possibility of the moral demand we make upon ourselves. Yet this further argument presupposes the first *Critique*'s argument that we cannot ground the principles of natural science themselves without at the same time revealing that their scope is limited to mere appearances.

Kant's bold attempt to resolve with one stroke two of the most pressing problems of modern philosophy has seldom been accepted by his successors without qualification. Some feel that Kant's identification of the basic principles of science with the fundamental principles of human understanding itself betrays too much confidence in the specifically Newtonian mechanistic physics that prevailed at his time, leaving too little room for subsequent scientific developments, such as the theory of general relativity and quantum mechanics. Others have felt that Kant's reduction of the laws of science to the laws of human thought is not an adequate account of the truly objective validity of science. Few have felt comfortable with the idea that the possibility of freedom could be defended by placing the real arena of human decision making behind a veil of ignorance, and many have felt that the idea that human freedom is our ultimate value but that it can be realized only through adherence to law is a strange and paradoxical one. Yet at the same time, broad elements of Kant's philosophy have become indispensable and therefore often almost invisible assumptions of the modern frame of mind. No modern thinker can believe that the human mind is merely a passive recorder of external fact, or a "mirror of nature."<sup>19</sup> But although many hold that since we have no way of stepping outside the human point of view, it may not be as easy as Kant thought to separate out our subjective contributions to the constitution of nature, yet every modern philosophy holds in some form or other the Kantian thesis that human beings make an active contribution to their knowledge. And although few defend human freedom through a rigid Kantian distinction between phenomenal appearance and noumenal reality, even fewer have thought that the assumption of causal determinism in science precludes conceiving of ourselves as agents who make decisions according to what seem to us to be the most rational principles of value. Thus many have accepted in some form the Kantian idea that there is a fundamental dif-

## Introduction

ference between the standpoints of the actor and the spectator,<sup>20</sup> and that this difference is crucial to the solution of the problem of free will. Even those who reject Kant's solutions to the problems of grounding natural science and making sense of our moral agency must solve these problems and find a way to avoid what they find objectionable in Kant's solution to them. In this way, all modern thinkers are children of Kant, whether they are happy or bitter about their paternity.

### III. THE EVOLUTION OF THE CRITIQUE

The *Critique of Pure Reason* has often been represented as the product of a violent revolution in Kant's thought that took place around 1772 – a midlife crisis in which the forty eight year old thinker rejected his previous adherence to the Leibnizian Wolffian philosophy, the systematic philosophy that Christian Wolff (1679–1754) had created out of the brilliant fragments that were all that was then known of the philosophy of Gottfried Wilhelm Leibniz (1646–1716) and that had become the dominant philosophy in enlightened German universities after the 1720s. Kant himself gave rise to this legend with several of his own remarks, above all his comment in the introduction to his *Prolegomena to Any Future Metaphysics* – the short work that Kant published in 1783 to try to overcome the initially indifferent or hostile reception of the *Critique* – that "it was the recollection of David Hume that many years ago first interrupted my dogmatic slumber and gave an entirely different direction to my investigations in the field of speculative philosophy."<sup>21</sup> There were certainly major changes in Kant's thought both before and after the publication of his inaugural dissertation, *De mundi sensibilis atque intelligibilis forma et principiis* (On the Form and Principles of the Sensible and Intelligible World) in 1770, the last publication preceding the years of intense but unpublished work leading up to the publication of the *Critique* in 1781. Nevertheless, Kant has misled those who have supposed that all his work in the years preceding this point was slumbering in Wolffian dogmatism, and that he awoke from this slumber only through some sudden recollection of the skepticism of David Hume (1711–1776).

In fact, Kant had been chipping away at fundamental tenets of the Leibnizian-Wolffian synthesis at least since the publication of his first exclusively philosophical work, his M.A. thesis *Principiorum primorum cognitionis metaphysicae nova dilucidatio* (A New Elucidation of the First Principles of Metaphysical Cognition) in 1755. There were certainly major developments in the content of Kant's philosophical views in the period around 1769–70 leading to the publication of the inaugural dissertation, and then further developments in Kant's doctrines and his