

Cambridge University Press
 978-0-521-65729-7 - Critique of Pure Reason
 Immanuel Kant
 Excerpt
[More information](#)

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.**"¹ 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).² 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.³ Another position Kant rejected was *empiricism*, which understood the “way of ideas” described in John Locke's (1632–1704) *Essay concerning Human Understanding* (1690) as grounding knowledge solely on ideas acquired in the course of individual experience. Yet another philosophical stance Kant encountered was what he called *indifferentism*, 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.⁴

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 Lockean 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,⁵ 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;⁶ 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;⁷ 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.⁸ 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 (1714–1791),⁹ 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*:¹⁰ 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.”¹¹ 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.¹² 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 (A 19/B 33), 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.¹³ 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 (A 21–2/B 36), 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.¹⁴ 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 (A 22–5/B 37–41, A 30–2/B 46–9).

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,¹⁵ 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 (A 26–30/B 42–5, A 32–48/B 49–73). 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” (A 66–81/B 91–116), which he named in the second edition of the *Critique* the “metaphysical deduction” of the categories (B 159), 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.¹⁶

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.¹⁷

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