Immanuel Kant’s *Critique of Pure Reason* is one of the seminal and monumental works in the history of Western philosophy.\(^1\) 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.

### 1 The Argument of the *Critique*

**The strategy of the *Critique.*** In the conclusion to his second critique, the *Critique of Practical Reason* (CPrR) 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” (CPrR 5: 161). 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 establish the basic principles of scientific knowledge of the world, thereby explaining our knowledge of the “starry heavens.” 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 first *Critique* is complex, however, because Kant argues on several fronts. Like Descartes, Locke, and Hume, Kant 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 (1714–62). Their position, which Kant called “dogmatism,” was compared in the preface to the

---

1. This text is an abridged and updated version of the Editor’s Introduction to our translation of the *Critique of Pure Reason*, published by Cambridge University Press in 1998. All citations to that edition are by the first (“A”) and second (“B”) edition pagination of Kant’s original text, reproduced in our translation. All other works of Kant are quoted from the volumes of the Cambridge edition listed at the end of this work, and passages are located by the volume and page number of the *Akademie* edition, also listed at the end, which are reproduced in the margins of all volumes of the Cambridge edition.
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 equally dangerous. The first of these is skepticism, the position Kant took David Hume (1711–76) to advocate. Another was the “way of ideas” described in John Locke’s (1632–1704) Essay Concerning Human Understanding (1690), which grounded knowledge solely on ideas acquired in the course of individual experience. Yet another philosophical stance was what Kant called indifferentism; it did not reject metaphysical assertions themselves, but scorned the attempt to argue for them systematically and rigorously. Here he had in mind a number of popular philosophers who agreed with dogmatists on metaphysical issues such as the existence of God and the immortality of the soul, but were unconvinced by the scholastic subtlety of their propositions and proofs. They held that the metaphysical beliefs we need are given simply through “healthy understanding” or common sense.

While Kant attempted to criticize and limit the scope of traditional rationalist metaphysics, he also sought to defend against the possibility of a priori knowledge, knowledge originating independently of experience, against the empiricist point that no particular experience (or a posteriori knowledge) could justify claims to universal and necessary validity. Just as Kant compared dogmatic metaphysicians to defenders of despotism, so he likened skeptics to nomads who abhor any form of permanent civil society 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, limited to what is required for ordinary experience and its extension by natural science. Thus Kant had to fight a war on several different fronts, in which he had to establish against both dogmatists and empiricists that many metaphysical questions are unanswerable; but he had to defend against both empiricists and skeptics parts of the very positions he was attacking, such as the possibility of a priori cognition of the fundamental principles required for natural science. And although he wanted to prove to the indifferentists that a science of metaphysics is important, he also wanted to embrace part of their position as well. For in regard to some insoluble metaphysical questions, Kant thought that we must defend a kind of commonsense belief in God, freedom, and immortality because morality has a vital stake in them.
The structure of the Critique. This complex program helps explain the enormous complexity of the structure and argument of the *Critique of Pure Reason*. To many readers, the “architectonic” of the *Critique* has been a barrier to understanding it, but a brief overview of the main divisions can illuminate its contents. Despite the profound originality of its contents, 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.” Such a distinction was common in German logic textbooks; in Kant’s hands, it becomes a distinction between his fundamental exposition of *a priori* cognition and its limits, in the “Doctrine of Elements,” and subsequent reflections on its methodological implications. The “Doctrine of Method” 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” is divided into two main (although very disproportionately sized) parts: the “Transcendental Aesthetic” and the “Transcendental Logic.” The former considers the *a priori* contributions of the fundamental forms of our *sensibility*, namely space and time, to our knowledge; the latter considers the *a priori* contributions, both genuine and spurious, to our knowledge made by the human intellect. This division is derived from Baumgarten, who introduced “aesthetics” as the title for the science of lower or sensitive cognition in contrast to “logic” as the science of higher or conceptual cognition. At the time he wrote 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 (A21/B35–6). After briefly distinguishing between “general logic” and “transcendental logic” – the former being the basic science of the forms of thought regardless of its object and the latter the science of the basic forms for the thought of objects (A50–7/ B74–82) – Kant then splits the “Transcendental Logic” into two main divisions, the “Transcendental Analytic” and the “Transcendental Dialectic.” This distinction derives from a sixteenth-century Aristotelian distinction between the logic of truth and the logic of probability, a distinction employed in eighteenth-century Germany by the Jena professor Joachim Georg Darjes (1714–92). Kant uses it instead to distinguish between the positive

---

contributions of the understanding, working in cooperation with sensibility (as
described in the “Transcendental Analytic”), and the false and deceptive
attempts of reason working independently of sensibility to provide metaphys-
ical insight into objects beyond possible experience (the “Transcendental
Dialectic”). The “Transcendental Analytic” is then further divided into two
books: the “Analytic of Concepts” and the “Analytic of Principles.” The former
argues for the universal and necessary validity of the pure concepts of the
understanding, or the categories, such as the concepts of substance and caus-
ation; the latter 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 the
Concepts of Pure Reason” and “On the Dialectical Inferences of Pure Reason.” In
the first, Kant explains how pure reason generates ideas of metaphysical
entities such as the soul, the world as a whole, and God; the second explores
critically the attempts to prove the reality of those ideas by extending beyond their
proper limits the patterns of inference which are valid only within the limits of
human sensibility. The division of the “Transcendental Analytic” into the
“Analytic of Concepts” and “Analytic of Principles,” then followed by the
“Dialectical Inferences of Pure Reason,” replicates the traditional division of
logic textbooks into three sections on concepts, judgments, and inferences. 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, but that
inferences of pure reason performed without respect to the limits of sensibility
give rise only to metaphysical illusion. The treatment of dialectical inferences is
in turn divided into three sections, “The Paralogisms of Pure Reason,” “The
Antinomy of Pure Reason,” and “The Ideal of Pure Reason.” In each section Kant
exposes metaphysically fallacious arguments about, respectively, the nature of the
soul, the world as a totality, and the existence of God. 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 the-
ology.” Kant replaces their “ontology” with the constructive doctrine of his
own “Transcendental Analytic” (see A247/B303), and then presents his criticism of
dogmatic metaphysics based on pure reason alone by demolishing the special
metaphysics of rational psychology, cosmology, and theology.

Kant divides the “Doctrine of Method” into four chapters, the “Discipline,”
the “Canon,” the “Architectonic,” and the “History of Pure Reason.” 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 that of philosophical argument, and offers important commentary on his own “critical” or “transcendental” method. In the “Canon of Pure Reason,” he prepares the way for his subsequent moral philosophy, contrasting the method of theoretical philosophy to that of practical philosophy. He also gives the first outline of the argument that runs through all three critiques, namely that although theoretical reason can never yield knowledge of God, freedom, and immortality, practical reason can justify rational belief or faith in such things. The last two parts of the “Doctrine of Method,” the “Architectonic of Pure Reason” and the “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. In the few pages of his “History of Pure Reason” Kant outlines the history of modern philosophy in terms of the distinction between rationalism and empiricism that we still use, while representing his own critical philosophy as the transcendence of both.

Within the organization of the Critique of Pure Reason as just presented, we now provide a brief résumé 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’s greatest influence has been exercised by Kant’s positive doctrine of the a priori elements of human knowledge. In the introduction, Kant argues that our mathematical, physical, and everyday knowledge of nature requires certain judgments that are “synthetic” rather than “analytic”; that is, they go beyond what can be known solely in virtue of the contents of the concepts involved in them – and yet these judgments are also knowable a priori, that is, independently of any particular experience. He entitles the question of how synthetic a priori judgments are possible the “general problem of pure reason” (B19), and proposes an entirely new science in order to answer it (A10–16/B24–30).

This new science Kant calls “transcendental” (A11/B25). It 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 even for the possibility of any cognition of objects. 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. Experience, however, is the product both of external objects affecting our sensibility and of the operation of our cognitive faculties in response to this effect (A1, B1). Kant’s claim is that we can have “pure” or a priori cognition of the contributions to experience that are made by
the operation of these faculties themselves rather than by the effect of external objects on us. Kant divides our cognitive capacities into **sensibility**: our receptivity to the effects of external objects acting on us, through which these objects are given to us in empirical intuition, and **understanding**: our active faculty for relating the data of intuition by thinking them under concepts and forming judgments about them (A19/B33).

**“Transcendental Aesthetic”: space and time; transcendental idealism.** Despite its brevity, 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 large proportion of the scholarly interpretation and controversy devoted to the *Critique* over the last two centuries. 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). He argues that space and time are pure forms of all intuition contributed by our own faculty of sensibility, and therefore that we can have *a priori* knowledge regarding them. Here Kant also attempts to resolve the debate about space and time 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 nonrelational properties inhering in spatiotemporally related things. 3 Kant’s new alternative 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. Only in this way, Kant argues, can we adequately account for our knowledge of the geometrical and mathematical properties of space and time as singular but infinite magnitudes.

Kant hereby rejects Leibniz’s account of space and time as mere relations abstracted from antecedently existing objects and accepts a Newtonian conception of absolute space and time, but only as the structure of human representation, not of the divine representation of reality (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 *transcendently ideal*, and so are the objects given in them. Although the precise meaning of this claim remains subject to debate, 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, from which Kant inferred 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/).

---

3. The correspondence between Leibniz and Newton’s spokesman Samuel Clarke was published in 1717, and in German in 1720.
B49–73). This is Kant’s transcendental idealism, which is employed throughout his critical philosophy in a variety of ways, both positively and negatively, and has been interpreted, attacked, and defended in a wide variety of ways over more than two centuries.

“Transcendental Analytic”: the metaphysical and transcendental deductions.

In the first part of the “Transcendental Analytic,” 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 experience. These twelve basic concepts, which Kant calls the categories, are fundamental concepts of an object in general or the fundamental general concepts of objects, that is, 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 in the second edition of the Critique Kant named the “metaphysical deduction” (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 those objects. There are four main logically significant aspects of judgments: their quantity, or the scope of their subject-terms; their quality, which concerns whether and how the predicate-term is affirmed or denied; their relation, or whether they assert a relation just between a subject and a predicate or between two or more subject-predicate judgments; and their modality, 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, that is, possibly, actually, or necessarily true. Corresponding to these twelve logical possibilities, Kant posits twelve fundamental categories for conceiving of the quantity, quality, relation, and modality of objects (A70/B95, A80/B106). Kant’s claims about the logical functions of judgment and the twelve corresponding categories have remained controversial ever since Kant first made them.

Kant’s even more ambitious claim is that these concepts apply universally and necessarily to the intuitions that are given in our experience, so that in some sense all of our experience (even of ourselves) is objective, that is, is subject to concepts that function as rules for the determinate relation of representations in experience (B137). (It must be kept in mind that for Kant “object” does not necessarily mean external object; rather, it will be key to Kant’s whole argument
that inner sense or internal experience is also part of the rule-governed world and in that sense objective). Kant argues for this thesis 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). In both versions of the Critique, Kant’s argument rests on the premise that any and all of our experiences can be ascribed to a single identical subject, through a mental act he calls the “transcendental unity of apperception.” Experience is possible only if the elements given in intuition are synthetically combined through this activity so as to present us with objects that are thought through the categories. The categories are therefore held to apply to objects not because the objects make the categories possible, but rather because combination according to the categories makes possible the representation of objects of experience. We do not abstract the categories from the experience of objects, but constitute the experience of objects by the use of the categories.

Principles of pure understanding. Even if the transcendental deduction succeeds in establishing that the categories do apply to all possible data for experience, it does so only abstractly and collectively. That is, it does not specify how each category applies necessarily to the objects given in experience, nor does it show that all of the categories must be applied to objects. To establish these further claims 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 some interpreters have thought sufficient to prove the objective validity of the categories independently of the prior transcendental deduction. In the third chapter Kant argues that because the categories have a determinate use only when applied to spatiotemporal data, they also have a determinate cognitive use only when applied to appearances (“phenomena”). By means of the categories things as they are in themselves (“noumena”) might be thought but they cannot be 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 the corresponding logical form or relation. In particular, Kant argues that each category must be associated with a temporal schema. 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 is that 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” (A144/B183). 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. The principles expressing the universal and necessary application of the categories to objects given in space and time are precisely 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 of Pure Understanding,” 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 applicability 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.
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. 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 that objects (more precisely, 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). Many interpreters consider the chapter on the Analogies of Experience the most important section of the Critique because it is supposed to supply the answer to Hume’s skeptical doubt about causality, while the third Analogy is the basis of Kant’s rejection of Leibniz’s denial of real interaction between substances. Both what the second Analogy is intended to prove and how the proof is supposed to proceed have been disputed almost as intensely as the philosophical question as to 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 justified theoretical 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 (4:288–94) that his “transcendental idealism” is a “critical” or “formal” idealism, which unlike traditional idealism does not deny the real existence of the objects distinct from ourselves that are represented as being in space and time.

In the third chapter of the “Analytic of Principles,” on phenomena and noumena, Kant emphasizes that because the categories must always be applied