

## Introduction

In the fifth and concluding section of the 1770 *Inaugural Dissertation*, Kant draws a distinction between mathematics and natural science, on the one hand, and metaphysics, on the other. He writes that while in mathematics and natural science ‘*use gives the method*’ (2:410), in ‘pure philosophy’ ‘*method precedes all science*’ (2:411). In metaphysics, a preliminary clarification of the proper ‘method’ is needed because we risk improperly using principles that belong to ‘sensibility’ to represent objects of the ‘understanding’, which are its subject matter. The *Critique of Pure Reason* gives up the idea that metaphysics should comprise cognitions of objects of the understanding that are not given through sensibility. However, Kant’s claim that a clarification of the ‘method’ of metaphysics should precede the actual science foreshadows the idea that metaphysics requires a propaedeutic ‘critique’ of reason.<sup>1</sup> Moreover, it is not the case that reference to a ‘critique’ completely substitutes talk of an inquiry focused on the ‘method’ of metaphysics. In fact, in two passages of the *Critique of Pure Reason*, Kant explicitly identifies his ‘critical’ investigation with such inquiry. In the 1787 Preface, he submits that the *Critique* ‘is a treatise on the method, not a system of the science itself’ (Bxxii). In a passage from the first edition that remains unaltered in the second, he labels the investigation he is pursuing a ‘doctrine of method’ (A82–3/B108–9).

But what could it mean to say that the *Critique* is a doctrine or treatise on method?<sup>2</sup> If we take a treatise on method to be an exposition of the procedures of investigation or argument that are appropriate in a particular discipline, it is difficult to understand Kant’s contention. The identification of such procedures is only partially pursued in the Discipline of Pure

<sup>1</sup> As will become clear, to say that the *Critique* is ‘propaedeutic’ is not to say that it comes before any part of metaphysics has been established.

<sup>2</sup> Scholars who have taken this statement seriously include Barale (1988), Tonelli (1994), La Rocca (2003: Ch. 6), Ferrarin (2015) and McQuillan (2016).

Reason, but this is insufficient to claim that the whole *Critique* has this aim. We find a similar problem if we focus on the second main part of the *Critique* in particular, the Transcendental Doctrine of Method: while the title suggests that it focuses on methodological issues, only the Discipline of Pure Reason fits this description. By contrast, it is not obvious why, for example, the Canon of Pure Reason or the History of Pure Reason are included in that part. Furthermore, an additional challenge comes from the fact that Kant characterizes both the *whole Critique* and one *part* of it as a ‘doctrine of method’, which is confusing.

The main aim of this book is to provide a satisfactory answer to these questions. In short, my overarching answer is that the claim that the *Critique* is a ‘doctrine’ or ‘treatise’ on method signals that it is the ‘doctrine of method’ of metaphysics. According to Kant, ‘doctrines of method’ have specific characteristics. It is by attending to these characteristics that we can solve the difficulties I have just mentioned. First of all, the principal task of a doctrine of method is *not* to identify procedures of investigation or argument that are appropriate within a particular science. Rather, their principal task is to show that a set of cognitions can be considered a science because it forms a ‘system’ with a certain unity, which Kant calls ‘architectonic’. In this respect, the *Critique* can be a ‘doctrine of method’ even though it only cursorily singles out procedures of investigation or argument that are adequate for a particular kind of investigation. Additionally, Kant submits that ‘doctrines of method’ usually come ‘at the end’ of a science, ‘because only then am I acquainted with the nature of the science’ (24:795). In my view, this explains why both a part of the *Critique* is titled ‘Doctrine of Method’ and the entire book can be considered such a doctrine. Insofar as a doctrine of method rests on the previous establishment of at least some parts of the science of which it is a doctrine of method, the *Critique of Pure Reason* requires that at least some parts of metaphysics be established in its ‘doctrine of elements’. Only in this way will it be able to perform its task as a doctrine of method. Since I claim that the *Critique of Pure Reason* is the doctrine of method of metaphysics and that, as such, it rests on the establishment of some parts of that science, it is here useful to clarify what I mean by metaphysics and what its ‘parts’ are, according to Kant.

## I Kant on Metaphysics in the *Critique of Pure Reason*

In the *Critique of Pure Reason*, Kant uses the term ‘metaphysics’ to refer to different things. For example, when he speaks of metaphysics as a ‘natural predisposition’ (B21–2), he has in mind reason’s tendency to ask questions

## Introduction

3

and advance claims about ‘unconditioned’ objects that lie beyond possible experience, such as God and the soul. However, he also speaks of metaphysics as *including* valid a priori cognitions of the ‘conditioned’ objects of possible experience, the validity of which is established in the *Critique* (Bxviii–xix). When I claim that the *Critique of Pure Reason* is the doctrine of method of metaphysics, I use the term ‘metaphysics’ in this second, more inclusive sense.

In order to gain an overview of the aims and structure of metaphysics, let us focus on Kant’s sketch of this science in the Architectonic of Pure Reason. There, metaphysics is seen as a part of ‘pure’ philosophy, which he describes as ‘cognition from pure reason’ (A840/B868). Roughly, this means that pure philosophy collects a priori cognitions, which Kant calls ‘philosophical’ in order to distinguish them from ‘mathematical’ a priori cognitions (A841/B869).<sup>3</sup> Pure philosophy is either propaedeutic, which is the critique of pure reason, or metaphysics, which comprises ‘the whole (true as well as apparent) philosophical cognition from pure reason in systematic interconnection’ (A841/B869).<sup>4</sup> Metaphysics, as the whole system of philosophical cognitions, is divided into the metaphysics of nature and the metaphysics of morals. The former identifies a priori concepts and principles for the theoretical cognition of objects. The latter singles out concepts and principles that determine ‘action and omission *a priori*’ (A841/B869). The metaphysics of nature is further divided into transcendental philosophy, which ‘considers only the understanding and reason itself in a system of all concepts and principles that are related to objects in general, without assuming objects that **would be given**’,<sup>5</sup> and the physiology of reason, which considers ‘nature’ as ‘the sum total of **given** objects’ (A845/B873). Finally, physiology is either ‘immanent’ or ‘transcendent’. ‘The former pertains to nature so far as its cognition can be applied in experience (*in concreto*), the latter to that connection of the objects of experience which surpasses all experience’ (A845–6/B873–4). Kant includes rational physics and rational psychology within immanent physiology and rational cosmology and rational theology within transcendent physiology (A846–7/B874–5).

There is much that remains obscure in this brief presentation of the aims and structure of metaphysics. Since our focus is the relationship

<sup>3</sup> Mathematical cognitions differ from philosophical cognitions because they rest on the ‘construction’ of concepts in intuition.

<sup>4</sup> In fact, Kant suggests that the term metaphysics can also be used to describe the whole of pure philosophy, including its propaedeutic part. This is clearly not the sense of ‘metaphysics’ I am using when I claim that the critique is the doctrine of method of this discipline.

<sup>5</sup> In Part II, I argue that transcendental philosophy also considers representations belonging to sensibility.

between the critique of pure reason and the projected science of metaphysics, we can single out three problems in particular. First, it is not clear how we can distinguish between the critique of pure reason, as the propaedeutic to metaphysics, and transcendental philosophy, as that part of metaphysics that identifies a priori concepts and principles that determine our cognition of objects in general. After all, it seems that in the *Critique* Kant dedicates much effort to identifying and establishing the validity of concepts and principles of this kind. Second, if the *Critique* is propaedeutic to the whole of metaphysics, as Kant's description suggests, in what sense is it also relevant to the metaphysics of morals? For example, it does not provide any justification for a priori principles of morality.<sup>6</sup> Finally, how should we take Kant's inclusion of rational physics, rational psychology, rational cosmology and rational theology within the system? Rational physics probably refers to what will become Kant's *Metaphysical Foundations of Natural Science*. In it, Kant identifies principles belonging to the special metaphysics of corporeal nature, which form the a priori part of a science whose objects are given empirically (see 4:469–70). This agrees with Kant's description of rational physics as included within 'immanent' physiology. At the time Kant was writing the first edition of the *Critique*, he still thought that an empirical psychology was possible and that it would similarly require an a priori part.<sup>7</sup> Therefore, Kant's reference to rational psychology as the second part of immanent physiology can be taken to refer to the 'metaphysical foundations' of psychology. But what should we make of Kant's inclusion of rational cosmology and rational theology in the picture? Are Kant's arguments in the Antinomy and the Ideal not designed to show that rational cosmology and rational theology cannot offer valid cognitions of objects and so cannot become sciences?

## 2 The Critique of Pure Reason, Transcendental Philosophy and the System of Metaphysics

I submit that the aforementioned problems dissolve when we read the *Critique of Pure Reason* as the doctrine of method of metaphysics. The

<sup>6</sup> Of course, Kant did provide this justification in the second *Critique*. However, the passage in the Architectonic does not explicitly maintain that the first *Critique* only partially accomplished the propaedeutic to metaphysics. This suggests, first, that at the time he was writing the A-version of the first *Critique* Kant viewed the first *Critique* as completing the critical task; second, it suggests that the *Critique's* relevance for moral metaphysics cannot lie in providing a 'foundation' for it.

<sup>7</sup> In a 1785 letter to Christian Gottfried Schütz, Kant still speaks of the 'metaphysical foundations' of psychology (10:406). The view that there can be such metaphysical foundations is rejected in the *Metaphysical Foundations of Natural Science*, however (4:471).

## Introduction

5

first problem I mentioned regarded the relationship between the critique of pure reason and transcendental philosophy. Even though Kant presents them as distinct in the *Architectonic*, much of what he does in the *Critique* seems to belong to transcendental philosophy. I have suggested that the *Critique*, as the doctrine of method of metaphysics, cannot perform its task if no part of metaphysics has already been established. But this gives us a tool for explaining the presence of arguments belonging to transcendental philosophy within the pages of the *Critique*. Put simply, the *Critique* contains elements of transcendental philosophy because they are instrumental to its role as the doctrine of method of metaphysics. This means that there are two disciplines that are established within the pages of the *Critique*: transcendental philosophy, as one part of metaphysics, and the critique of pure reason, as that discipline within the *Critique* that achieves the latter's aim as the doctrine of method of metaphysics.<sup>8</sup> In this respect, my approach turns the usual view on the relationship between these disciplines on its head. Typically, they are seen as fundamentally overlapping, or, alternatively, the critique of pure reason, as a propaedeutic, is thought to come before the establishment of any part of transcendental philosophy.

I characterize transcendental philosophy as that part of the metaphysics of nature that investigates a priori concepts for the cognition of objects that do not contain anything empirical (see A845/B873).<sup>9</sup> The part of transcendental philosophy that is established within the *Critique* takes into consideration not all of these concepts but only those that Kant calls 'root concepts' (*Stammbegriffe*) (see A14/B27–8),<sup>10</sup> which are concepts that lie at the basis of synthetic a priori claims. With respect to these concepts, transcendental philosophy has two main tasks. First, it identifies concepts that are candidates for meeting this characterization and determines their origin. Second, it examines their validity. The first task is performed by metaphysical deductions, the second by transcendental deductions. This approach to transcendental philosophy is innovative in at least two respects.

<sup>8</sup> Let me here add a terminological note. Throughout this book, I will use '*Critique of Pure Reason*' (in italics and with capitalization) to refer to Kant's book in its entirety. By contrast, I will use 'critique of pure reason' (in roman font and without capitalization) to refer to that discipline within the *Critique* that achieves the latter's aim as the doctrine of method of metaphysics.

<sup>9</sup> Admittedly, 'a priori concepts for the cognition of objects' is a vague formulation. I adopt it because if we consider the concepts studied by transcendental philosophy, they seem to be relevant to the cognition of objects in different ways. Some, like the categories, are constitutive of our cognition of objects. Others, like the ideas of reason used regulatively, are conditions for attaining some cognitions of objects but are not themselves constitutive of those cognitions.

<sup>10</sup> I here follow Pluhar's translation of the first *Critique* in translating *Stammbegriff* as 'root concept'.

It characterizes the investigation into the validity of a priori cognition as only being concerned with establishing *positive* results. That is, it is not the task of transcendental philosophy to set *limits* to our cognitions. Rather, it determines where and why an a priori cognition is valid. Additionally, I argue that the distinction between the two tasks of transcendental philosophy can be drawn in all main parts of the Transcendental Doctrine of Elements, not only in the Transcendental Analytic. There is a metaphysical and a transcendental deduction of space and time in the Aesthetic, a metaphysical and a transcendental deduction of the categories in the Analytic, and a metaphysical and a transcendental deduction of the ideas of reason in the Dialectic.

Let us now move to the second problem. In what sense is Kant's first *Critique* relevant to the metaphysics of morals? Clearly, it does not provide an analysis of or foundation for moral principles. According to my approach, the *Critique* is relevant to the metaphysics of morals because it shows that the latter can form a coherent part of metaphysics as a whole. This does not necessarily involve providing a direct justification of moral principles, but it brings us to a characterization of the second discipline established in the *Critique*, namely, the critique of pure reason as the doctrine of method of metaphysics. As such, the critique must show that metaphysics as a whole (comprising both its theoretical and its practical parts) can become a science because it can achieve 'architectonic unity'. In my account, 'architectonic unity', while being a condition of science, is different from mere systematicity. In order to attain architectonic unity, a science must realize what Kant calls its 'idea', which I take to be the *correct* description of the body of cognitions belonging to a science and the parts-whole relationships within it.

In order to legitimately attribute architectonic unity to a body of cognitions, the latter must at least meet two minimal conditions. First, it must possess *systematic coherence*. I take a body of cognitions to be systematically coherent when: (a) the cognitions belonging to it are interconnected in a way that involves relations of either logical implication, explanatory support or both; and (b) it does not contain contradictions. Second, it must be possible to view the body of cognitions as realizing the fundamental 'idea' of a science, where this idea must (i) define the fundamental object of that science and (ii) prescribe the ordering of the body of cognitions that form that science. The main obstacles to the systematic coherence of metaphysics are the disputes in special metaphysics that form what Kant calls the 'conflict of reason with itself'. Kant's solution to these disputes rests on drawing the negative consequences of the doctrines established in

*Introduction*

7

transcendental philosophy. Accordingly, the critique of pure reason must first of all show how the very way in which the validity of root concepts is proved by transcendental philosophy implies that their validity is limited, which in turn can put a stop to the metaphysical disputes in question.

Putting a stop to these disputes is insufficient for establishing that metaphysics can attain architectonic unity, however. The critique must also show that metaphysics can indeed be seen as realizing its fundamental 'idea'. This clarifies, first, why the main task of the critique of pure reason is *positive* and, second, why the critique is relevant to the metaphysics of morals. For Kant, an essential part of the proper 'idea' of metaphysics is constituted by cognitions belonging to the practical part of metaphysics. The latter cognitions appear to demand a commitment to propositions in special metaphysics that the negative part of the critique deems beyond the scope of possible cognition. In this respect, the critique must show that we can meet the demands of the practical part of metaphysics without endangering the agnosticism regarding objects of pure reason that is established by the negative part of the critique of pure reason. The negative task of the critique, which is concerned with establishing limits to the use of root concepts, is thus best seen as merely *subordinate* to its main positive aim, that is, establishing that metaphysics can meet the two minimal conditions of architectonic unity. If this is right, we can see that the main task of the critique of pure reason, at least as it is conceived in the first *Critique*, is not to provide a justification of certain synthetic a priori principles (which is a task that is already part of metaphysics), but rather to show that a body of cognitions, including those very synthetic a priori principles, can form a whole with a proper unity that bestows the status of science to metaphysics.

Let us move to the third problem identified in Section 1. Why does Kant include rational cosmology and rational theology in his sketch of the structure of metaphysics? An 'easy' way to put aside this worry is to stress that Kant's sketch contains 'true as well as apparent' (A841/B869) cognitions. Accordingly, one could suggest that metaphysics includes analyses of the concepts of the world and God, but that it takes into account that they only constitute 'illusions', which is coherent with the results of the Dialectic. In my view, this explanation is unsatisfactory. It is unable to clarify why those 'apparent' cognitions should be part of metaphysics. I believe that a better answer can be obtained when we take the perspective of the critique in its relationship to the practical part of metaphysics. I have suggested that one chief task of the critique is to show that we can meet the demands of the practical part of metaphysics, where these demands



concern commitments towards objects that customarily belonged to special metaphysics, such as God, the world, and the soul. Even though the commitments in question are ‘practically’ justified for Kant, since our justification rests on a moral argument, the propositions to which we commit ourselves are ‘theoretical’ because they are descriptive and do not concern ‘oughts’. In this sense, they belong to the ‘theoretical’ part of metaphysics, even though they are grounded in its practical part and do not constitute ‘cognitions’. I believe that we should read at least Kant’s inclusion of rational theology within his sketch of metaphysics in this sense.<sup>11</sup>

### 3 Kant and the ‘System’ of Metaphysics

Reading the *Critique of Pure Reason* as the doctrine of method of metaphysics puts metaphysical concerns at the core of Kant’s project.<sup>12</sup> But did Kant view the *Critique* as presenting his ultimate answer to at least some of these concerns, or did he take his investigation to be merely provisional? This question is not trivial, since many contemporaries of Kant, who took themselves to be furthering his approach, thought that the *Critique* only paved the way for a metaphysical investigation that ultimately needed a different and more fundamental foundation. Kant bitterly reacted to these attempts to ‘complete’ his plan. Famously, in his open letter on Fichte’s *Wissenschaftslehre*, published in 1799, he wrote:

I must remark here that the assumption that I have intended to publish only a propaedeutic to transcendental philosophy and not the actual system of this philosophy is incomprehensible to me. Such an intention could never have occurred to me, since I myself, in the *Critique of Pure Reason*, have lauded the completed whole of pure philosophy as the best indication of the truth of that work. (12:370–71, translation altered)<sup>13</sup>

Kant’s open letter was motivated by an anonymous review of Johann Gottlieb Bohle’s *Entwurf der Transcendental-Philosophie*, in which the author of the review invited Kant to express his thoughts on Fichte’s *Wissenschaftslehre*, a project that, according to the reviewer, was the

<sup>11</sup> Unlike Ypi (2021: 16, 168–72), I do not see Kant’s inclusion of this discipline as involving a step back into dogmatism. As far as rational cosmology is concerned, recall that freedom is introduced as a cosmological idea and that it is essential from the perspective of the practical part of metaphysics.

<sup>12</sup> An emphasis on the relevance of these concerns has characterized various recent interpretations of Kant’s work. For example, Jauernig (2021) provides a ‘metaphysical’ interpretation of transcendental idealism; Willaschek (2018) provides an account of what we have called ‘metaphysics as a natural predisposition’; De Boer (2020) focuses on Kant’s projected system of metaphysics and its relationship to the Wolffian tradition.

<sup>13</sup> I here follow De Boer’s (2020: 251–2) suggested translation of the second sentence.



*Introduction*

9

completion of the work that Kant initiated with his propaedeutic ‘critique’.<sup>14</sup> The idea that Kant’s critical investigation was ‘provisional’ and needed to be brought to completion was not only the reviewer’s or Fichte’s but was shared by many philosophers who regarded themselves as ‘Kantians’, including Jakub Sigismund Beck and Karl Leonhard Reinhold. The common idea behind these projects was that Kant’s system remained incomplete because it lacked a unique and single principle that could serve as the ultimate foundation of metaphysics. In their view, without such a principle, metaphysics could not attain the unity of a system.

We can understand Kant’s irritation with the claim that his position needed a further foundation. Still, his astonishment at the suggestion that he ‘intended to publish only a propaedeutic’ seems unwarranted. After all, he himself stressed at various points that the *Critique* was only a ‘preparation’ to metaphysics. Accordingly, many interpreters have taken Kant’s open letter to sharply contradict his earlier view. In contrast to this approach, Karin de Boer has recently argued that the open letter is coherent with Kant’s description of the *Critique* as a propaedeutic. She emphasizes that in this passage Kant is challenging *not* the contention that the system is incomplete but the contention that he did not *intend* to complete it (De Boer 2020: 251–2). In other words, he did not deny that the *Critique* was a propaedeutic but only that he was content with it and did not want to complete it himself.

While De Boer is certainly right regarding what the passage literally says, I believe that Kant’s irritation was caused by a misrepresentation not only of what he *intended* to achieve, but also of what he took himself to have *already* achieved. This does not mean that we cannot make sense of his astonishment. I have already mentioned that the philosophers to whom Kant was responding viewed Kant’s position as incomplete because it lacked a proper foundation. This also applies to the analyses of ‘root concepts’ that, in my view, constitute the part of transcendental philosophy that is established in the *Critique of Pure Reason*. It is true that Kant did view transcendental philosophy as it is presented in the *Critique* as incomplete, but not because it lacked a proper ‘foundation’. Rather, what it still lacked was an identification of the ‘derivative’ concepts that rested on the root concepts singled out in the *Critique*. If we keep this in mind, we can understand Kant’s contentions in the passage without assuming any radical break with his earlier view. Kant took himself to have already established some *parts* of metaphysics within the pages of the *Critique*.

<sup>14</sup> The review appeared in the *Literatur-Zeitung* on Friday, 11 January 1799.

These parts correspond to *parts* of transcendental philosophy, which are not in need of a further foundation. This is compatible with viewing the *Critique* as playing a propaedeutic role with respect to metaphysics as a whole, which Kant clearly did not think he had ‘completed’ in the *Critique of Pure Reason*. Therefore, we can make perfect sense of Kant’s open letter as confirming that the critique of pure reason rests on the previous establishment of some parts of transcendental philosophy.

#### 4 The ‘Method’ of the *Critique of Pure Reason*

Reading the *Critique* as the doctrine of method of metaphysics also has consequences for investigations aimed at determining the ‘method’ that Kant follows in the *Critique of Pure Reason*. Here, ‘method’ means a particular procedure of argument. These investigations often revolve around what a ‘transcendental argument’, a ‘transcendental proof’, or a ‘transcendental deduction’ is for Kant (see for example Strawson 1966; Henrich 1969; Henrich 1989; Carl 1992; Engstrom 1994; Cassam 1987; Ameriks 1978; Ameriks 2003; Hatfield 2003; G. Bird 2006b; Callanan 2006; Stapleford 2008; Moore 2010; Pereboom 2019). The purpose of these discussions is mainly to evaluate the aims, structure and validity of famous arguments from the *Critique of Pure Reason*, such as the transcendental deduction of the categories, the refutation of idealism and the second analogy, where these arguments are taken to be paradigmatic of Kant’s method. Alternatively, some interpreters have focused on the notion of ‘transcendental reflection’ that Kant introduces in the Amphiboly of the Concepts of Reflection as a key to unravelling what is distinctive about Kant’s argumentative strategy in the *Critique* (Longuenesse 1998: Chs. 5–6; Leitner 1994; Smit 1999).<sup>15</sup> The task of transcendental reflection is to provide a ‘transcendental topic’ in which a concept is assigned its appropriate ‘transcendental place’ either in the understanding or in sensibility, according to the faculty through which the object of that concept is given to us (A268–9/324–5).<sup>16</sup>

If my suggestion that there are actually two disciplines established within the pages of the *Critique* is right, it seems to follow that the attempt to provide a *unitary* account of its method is doomed from the start, because we need at least to distinguish between the ‘method’ of transcendental

<sup>15</sup> One problem with this approach is that Kant’s position in the Amphiboly resembles his perspective in the *Inaugural Dissertation*. Accordingly, it is unclear whether the Amphiboly is completely coherent with Kant’s changed view in the *Critique*. See Willaschek (1998: 341–2).

<sup>16</sup> On transcendental reflection, see also De Boer (2020: Ch. 7), Merritt (2015) and Gava (2019a).