

## Kant on the Necessity of Metaphysics<sup>1</sup>

Marcus Willaschek, Frankfurt / M.

(To appear in: Proceedings of the 10. International Kant-Congress, Berlin: de Gruyter 2006)

“Human reason has this peculiar fate that in one species of its knowledge it is burdened by questions which, as prescribed by the very nature of reason itself, it is not able to ignore, but which, as transcending all its powers, it is also not able to answer” (KrV A1). This tragic constitution of human reason, as Kant points out some lines later, gives rise to “endless controversies” on the “battle-field” known as “metaphysics.” The famous first sentence of Kant’s Preface to the A-edition of the *Critique of Pure Reason* thus contains two claims that are central to Kant’s conception of metaphysics: First, metaphysical questions cannot be ignored since they are “prescribed by the very nature of reason itself”; in other words, they arise necessarily out of the very structure of rational thinking as such. And second, metaphysical questions cannot be answered (answered, that is, in a rationally responsible way), since they transcend the reach of human reason. Later in the first Critique, in the “Introduction to the Transcendental Dialectic,” both claims are radicalized as follows: First, the structure of rational thinking as such gives rise not only to metaphysical *questions*, but also to particular *answers* to these questions. This I will call the *necessity of metaphysics* thesis. The second claim then becomes what I will call the *necessity of metaphysical illusion* thesis: Since metaphysical questions cannot be answered in a responsible way, the answers that come naturally to human reason are apt to deceive us (unless deception is prevented by a critique of pure reason).<sup>2</sup> To these two claims, Kant later adds a third, which is also elaborated in the second and third Critiques: Metaphysical theses concerning God, freedom, and immortality can indeed be rationally upheld, not as metaphysical knowledge, but as “postulates of pure practical reason.” This third thesis I will call the *practical transformation of metaphysics*.

The first two theses and the arguments Kant offers for them form a profound and subtle diagnosis of traditional metaphysics, one that differs from earlier criticisms of metaphysics in that it locates the source of both metaphysics and its failure in the structure of rational thinking itself. The third claim, by contrast opens up an entirely new perspective not

---

<sup>1</sup> In writing the final version of this text, I have been greatly helped by comments on an earlier version from Patricia Kitcher, Eckart Förster, and Michelle Grier; I am aware, though, that I have not been able to do full justice to all of their suggestions. Thanks also to Shannon Hoff who revised my English. – Translations from the *Critique of Pure Reason* are Norman Kemp-Smith’s, from the *Critique of Practical Reason* Lewis White Beck’s, with minor revisions.

<sup>2</sup> This point is hinted at already in the second paragraph of the Preface A.

only on metaphysics, but also on the concept of rationality and thus on what it means to be a rational being. The three theses of the *necessity of metaphysics*, the *necessity of metaphysical illusion* and the *practical transformation of metaphysics* together make up a revolutionary conception of metaphysics whose philosophical import has still not been fully explored.

In what follows, I want to contribute to such an exploration. In the limited space of this essay, however, I will have to limit my focus in various ways. First, I will devote most of my attention to Kant's arguments for the necessity of metaphysics and its practical transformation, at the cost of his conception of metaphysical illusion. Second, I will neither discuss the status of the three disciplines of traditional "metaphysica specialis" (rational psychology, cosmology, and theology) nor prospects of metaphysics "as science", that is, metaphysics as part of Kant's project of a "transcendental philosophy" as developed in the *Metaphysical Foundations of Natural Science* and the *Metaphysics of Morals*. Rather, my discussion will be restricted to what Kant calls metaphysics as a "natural disposition" (3: 41; cf. 4: 362ff.). And finally, in discussion the necessity of metaphysics thesis, I will concentrate exclusively on Kant's critical exposition in Section Two of the "Introduction the Transcendental Dialectic", without considering his closely related doctrine of transcendental ideas (A312/B368 ff.) and of the "regulative" use of transcendental ideas and principles defended in the "Appendix to the Transcendental Dialectic" (A642/B670 ff.) and the *Critique of Judgement*. I hope, though, that these deplorable limitations of scope will not affect the main line of argument that I want to develop in this paper: I will argue that Kant indeed discovers a source of metaphysical thinking that lies in reason itself, as it is traditionally conceived. Its central feature is the iterative structure of reason-giving and explanation: That is, if "Why A?" is a good question and "Because of B" is a good answer to that question, then "Why B?" is a good question, too, that needs to be answered if reason is to be satisfied. It is this iterative structure that takes us, in Kant's words, from the conditioned through the complete series of conditions to the unconditioned. Contrary to what Kant assumes, however, I will argue that this conception of reason, although characteristic of western philosophy since the days of Plato and Aristotle, is merely optional. This becomes apparent, somewhat ironically, through Kant's own practical transformation of metaphysics, since Kant's argument for the "postulates of pure practical reason" employs a conception of reason that breaks with the traditional conception of rationality. This means that Kant's claims about the *necessity of metaphysics* and the *necessity of metaphysical illusion* turn out to contain no critique of human reason as such, but rather a critique of a particular *conception* of reason that

is characteristic of western metaphysics. – But let me begin with the question of why, according to Kant, metaphysics is necessary.

### 1. The Necessity of Metaphysics

According to Kant, metaphysics is necessary in that it arises from “the very nature of reason itself” (A1). Owing to the ambiguity of the term “nature” (cf. A418/B446 fn), this can either mean that metaphysics belongs to reason “essentially” or that it belongs to reason, considered as an innate and thus “natural” capacity of human beings. Kant subscribes to the claim under both interpretations. According to Kant, metaphysics exists in all human beings as a “natural disposition” (3: 41). This is so because reason (as a disposition) is part of human nature and “from the nature of universal human reason [...] questions arise which pure reason propounds to itself, and which it is impelled by its own need to answer as best it can” (3: 41). Thus, metaphysics comes natural to human beings because it belongs to the “essence” of reason to lead us into metaphysical thinking. The latter claim we might paraphrase as follows: Because of the very features that characterize a kind of thinking as rational, rational thinking leads to metaphysical questions and to metaphysical answers to them.

According to Kant’s understanding of the term, a claim is *metaphysical* if it concerns “the unconditioned” (cf. B XX) – that is, something which is as it is independently of, or unconditioned by, anything else. Kant considers three kinds of objects that qualify as unconditioned: the *soul* as the “absolute (unconditioned) unity of the thinking subject”, the *world* as “absolute unity of the series of conditions of appearance,” and finally *God* as “the absolute unity of the condition of all objects of thought” (A334/B391). The notions of soul, world, and God, according to Kant, are “transcendental ideas” – ideas, that is, which arise necessarily from human reason, but whose objects cannot be given in a possible experience. Thus must the necessity of metaphysics consist in the fact that rational thinking as such inevitably leads to claims about the unconditioned: about immortal souls, the world as a whole, and God.

Kant’s central argument for the necessity of metaphysics roughly runs as follows: Reason necessarily relies on the principle that, if something *conditioned* is given, then the complete series of its conditions must also be given. Now the series of conditions is either finite or infinite. If it is finite, then there must be at least one condition that does not depend on any further conditions. If it is infinite, then the series of conditions as a whole does not depend on any other condition. Hence, in both cases, reason leads us from the existence of something that is conditioned to the conclusion that there must be something unconditioned:

If the conditioned is given, then so is the unconditioned. Let's call this the "conditioned-unconditioned" principle. Although Kant calls this principle "the highest principle of pure reason" (A308/B365), and although the entire argument of the Transcendental Dialectic depends on it, it has received comparatively little attention from readers and commentators.<sup>3</sup> The reason for this cannot be that the principle and Kant's arguments for it present no problems. Quite to the contrary, they raise a number of tantalizing questions, both exegetical and philosophical, only some of which I can address here. Although the argument is repeated, with significant variations, in several places,<sup>4</sup> I will focus exclusively on Kant's official argument for the conditioned-unconditioned principle, presented in Part II of the "Introduction to the Transcendental Dialectic."

There, Kant begins as follows: "All our knowledge starts with the senses, proceeds from thence to understanding, and ends with reason, beyond which there is no higher faculty to be found in us for elaborating the matter of intuition and bringing it under the highest unity of thought" (A298f/B355). The idea that reason must bring about the highest unity of thought will turn out to be a key element in Kant's argument for the *necessity of metaphysics*, and hence it would be somewhat disappointing if Kant were to presuppose it by simply defining reason as the faculty that brings about such a unity. Kant, of course, is aware of this, and goes on to define reason as the "faculty of principles" (A299/B356). This definition is meant to be neutral with respect to Kant's distinction between the merely formal or logical employment of reason, and its real or transcendental employment. In its logical employment, reason is the faculty of mediate inference that allows us to deduce conclusions from two or more premises (A299/B355). In its real or transcendental employment, however, reason is a source of concepts and principles that is independent from other sources such as intuition and understanding (*ibid.*).<sup>5</sup>

Now both of these employments of reason fall under the definition of a "faculty of principles," because, as Kant points out, the term "principle" is ambivalent. On the one hand, it can mean any universal proposition that is used as a major premise in a syllogism. "Thus every syllogism [and hence every logical employment of reason] is a mode of deducing knowledge from a principle" (A300/B357). On the other hand, a principle is a piece of "synthetic knowledge from concepts" (A301/B357f.). Only these latter principles, for Kant,

---

<sup>3</sup> An important exception is Michelle Grier, *Kant's Doctrine of Transcendental Illusion*, Cambridge 2001, Ch. 4.

<sup>4</sup> In Book 1 of the "Transcendental Dialectic" (A330/B386-A332/B389) and in Sections 1 and 7 of the "Antinomy"-Chapter (A408/B435-A411/B438; A497/B525-A501/B529).

<sup>5</sup> Although terminology may suggest something else, Kant's explicit definitions make it clear that the "logical" use of reason belongs with his conception of "general logic," while the "real" use (in spite of its being contrasted with the logical use) belongs with his conception of a "transcendental logic" (cf. A55/B79 ff.).

can be called principles “without qualification,” whereas all other universal propositions may be called “comparative principles” (A301/B358). If we consider Kant’s definition of “synthetic judgements” as judgements in which the predicate-concept B lies entirely outside of the subject-concept A (A6/B10), the definition of principles as “synthetic knowledge from concepts” seems to be contradictory. Surely, Kant does not want to rule out the existence of principles of pure reason simply by definition. What Kant means by “knowledge from concepts,” however, is simply “knowledge from reason alone.” Kant gives no examples of (alleged) knowledge of this kind, but presumably he is thinking of something like Leibniz’ “principle of sufficient reason” or the “principle of complete determination” Kant discusses in the “Transcendental Ideal” (A573/B601 ff.). These principles are “a priori without qualification” (cf. B24) insofar as they do not rely on the possibility of corresponding experience; and they are synthetic insofar as they at least presume to convey substantial knowledge about the world. If there were such a principle, it would be what Kant later calls “a transcendental principle of reason,” which “would make the systematic unity necessary, not only subjectively and logically, as method, but objectively also” (A648/B676).

After introducing the distinctions between the immediate inferences of the understanding and the mediate inferences of reason, and between categorical, hypothetical and disjunctive syllogisms, which will not concern us here, Kant proceeds by tracing reason’s preoccupation with systematic unity to the very structure of syllogistic reasoning: “in inference reason endeavours to reduce the varied and manifold knowledge obtained through the understanding to the smallest number of principles (universal conditions) and thereby to achieve in it the highest possible unity.” (A305/B361) Kant’s argument for this claim begins with the observation that we often wonder whether something we already know follows as a conclusion from other, more general knowledge. Consider an example Kant uses later in a similar context: the proposition that all bodies are alterable (cf. A330/B387).<sup>6</sup> Every general proposition consists of a “condition” and an “assertion” (cf. 9: 121); in our example, “being a body” is the condition and “being alterable” is the assertion. We now “look in the understanding for the assertion of this conclusion to discover whether it is not there found to stand under certain conditions according to a universal rule” (A304/B361). So what we look for is a universal proposition with the same assertion (“being alterable”), but with a different condition, such as “everything composite is alterable.” “If I find such a condition, and if the

---

<sup>6</sup> In A330/B387 Kant uses the example differently in that he does not assume that we already know that the conclusion is true so that we can come to know its truth only by deriving it from more general premises. Here, in A304/B361, Kant assumes that the conclusion is known already “through the understanding,” as the final sentence of the section makes clear.

object of the conclusion can be subsumed under the given condition, then the conclusion is deduced from the rule, *which is also valid for other objects of knowledge*” (A304f./B361). So we subsume “being a body” under the condition of the rule that everything composite is alterable, and thus arrive at the minor premise “all bodies are composite,” which allows us to deduce our original judgement as a conclusion. In this way, we have subsumed a more particular piece of knowledge under a more general one, and have thus taken a step towards unifying our knowledge of nature.

The task Kant here ascribes to reason is that of giving a particular kind of *explanation*. It concerns universal propositions such as “all bodies are alterable,” which are already considered to be part of our knowledge since they are authorized by experience and the understanding. But even if we accept it as a fact that all bodies are alterable, we may still wonder *why* this is the case. As Kant conceives of it, the task of reason consists in answering this kind of question. Since the question is not which universal propositions hold, but *why* they hold, the way to answer them consists in subsuming the proposition in question under another universal proposition, which, as Kant puts it, “*is also valid for other objects of knowledge*.” So the aim is not only to derive more particular from more general knowledge, but also to relate a particular proposition to others in an illuminating fashion. If, for example, we find that all composites are alterable, we are not just able to subsume the proposition “all bodies are alterable” under a more general rule. If we consider for instance that gases are composites, too, we also see that the reason why bodies are alterable is the same as the reason why gases are alterable—namely, that both are composites. By pointing out the rational connections among previously unconnected pieces of knowledge, we develop a particular kind of explanation, which I shall call *inferential* explanation.

Now Kant’s next question is whether, in its search for inferential explanation, reason is limited to principles that are supplied by the understanding (principles restricted to the realm of possible experience), or whether “reason in itself, that is pure reason, contains a priori synthetic principles and rules” (A306/B363). These would not be merely comparative principles, but transcendental principles of pure reason.<sup>7</sup> Kant first points out that the inferences of reason do not refer to intuitions and thus directly to individual objects, but only

---

<sup>7</sup> In analogy to the derivation of categories and principles of the understanding from the logical forms in judgments, Kant claims that the “formal and logical procedure of reason in syllogisms gives us sufficient guidance as to the ground on which the transcendental principle of pure reason in its synthetic knowledge will rest” (A306/B363). This is important with respect to the necessity of metaphysics thesis, since it means that metaphysical principles are meant to be already implicit in the logical employment of reason and to grow out of it naturally.

as mediated through concepts and the general judgements of the understanding.<sup>8</sup> He then proceeds to derive what he calls “the principle peculiar to reason in general, in its logical employment” in a complex passage that I will quote in full:

“[R]eason, in its logical employment, seeks to discover the universal condition of its judgment (of the conclusion), and the syllogism is itself nothing but a judgment made by means of the subsumption of its condition under a universal rule (the major premise). Now since this rule is itself subject to the same requirement of reason, and the condition of the condition must therefore be sought (by means of a prosyllogism) whenever practicable, obviously the principle peculiar to reason in general, in its logical employment, is: -- to find for the conditioned knowledge obtained through the understanding the unconditioned whereby its unity is brought to completion” (A307/B364).

For several reasons, the last sentence of this passage is difficult to understand. First, it is not clear why Kant speaks of the peculiar principle of reason in its *logical* employment. Since the logical employment consists in drawing inferences, one might have expected another principle such as the one Kant in his *Logic* calls “the highest principle of all inferences of reason”, namely “What stands under the condition of a rule also stands under the rule itself” (9: 120). But obviously, in the passage under discussion Kant is not interested in the direction of inference “downwards” from premises to conclusions, but rather “upwards” from the conditioned to the condition. However, he gives no indication why the move upwards should belong to the *logical* employment of reason. A possible reason might be that in both its upward and its downward employment, reason is abstracting from content and considering only whether the conclusion follows logically from the premises. Secondly, while in the sentence before, Kant had used the term “condition” in the logical sense of “subject term in a general proposition”, so that the search for “the condition of the condition” is the search for judgements containing ever more general concepts in subject-place, Kant now speaks of “conditioned knowledge”. Presumably, knowledge is conditioned if it is based on, or derived from, some other knowledge. Since the conditioned knowledge is said to be “knowledge of the understanding”, one might assume that it is conditioned by the possibility of experience.

---

<sup>8</sup> Kant’s point here is related to his remark on the table of judgements that “in the employment of judgements in inferences of reason, singular judgements may be treated as universal” (A71/B96). The reason for this is that syllogisms as traditionally conceived cannot contain singular, but only general propositions of the form “Some A are B” or “All A are B”.

In a later passage, by contrast, Kant calls the premises in a syllogism “conditions” for the knowledge contained in the conclusion (A331/B387), which suggests that “conditioned knowledge” is knowledge conditioned by premises. Both readings of the passage under discussion make sense and I will have to leave undecided which one is correct. Third, the principle directs us to find “to the conditioned knowledge of the understanding the unconditioned” (“zu dem bedingten Erkenntnis des Verstandes das Unbedingte zu finden”). Now if we ignore that in the German original the “Unbedingte” is written with a capital U, one might take Kant to say that we must find unconditioned *knowledge*, which sounds just right: Reason climbs up the ladder of conditionally known premises until it arrives at a premise which is unconditioned in the sense that it need not, and cannot, be derived from any higher premise. In the section on Transcendental Ideas, Kant himself put things more or less that way (A332/B389 f.). But since Kant writes “das Unbedingte” with a capital U, what he does say is that reason, in its logical employment, must find “the unconditioned”, which suggest some object rather than a piece of knowledge. In fact, nowhere in the first Critique Kant calls knowledge “unconditioned.” It seems that, although Kant speaks of “conditioned knowledge,” the term “unconditioned” is reserved for the objects of transcendental ideas. But then, it is strange that reason in its *logical* employment should have to find the unconditioned, since in its logical employment reason, as Kant says, “abstracts from all content of knowledge” (A299/B355).

Setting these exegetical questions aside, the central philosophical question seems to be why, according to Kant, reason’s search for the “condition of the condition” should take us all the way to something unconditioned. The answer suggested in the final words of the cited passage (“whereby its unity is brought to completion”) is that only in this way can we achieve the particular kind of unity at which reason aims. Kant’s argument for this claim starts out from the feature of the logical employment of reason he previously introduced: if a judgement is supplied by the understanding, reason seeks premises from which that judgement can be derived. As Kant now points out, this leads to an iterative process, since each judgement that serves as a premise may be equally considered as a potential conclusion to be derived from higher, more abstract principles. Thus, once we have derived the judgement “All bodies are alterable” from the “rule,” “Everything composite is alterable,” then reason must look for an even more general law from which it can derive the “rule” that everything composite is alterable. Reason, of course, may fail. We may simply not find a more general law. Kant admits this possibility by saying that “the condition of the condition must therefore be sought [...] *whenever practicable*” (“wann immer es angeht”). But we must at least *look* for a



“condition of the condition,” that is, we must search after ever more general principles until we reach something “unconditioned.” Rational beings, according to Kant, not only try to systematize their knowledge as far as possible; they also seek a systematic unity that allows them to subsume *all* the laws of nature under a finite set of principles (or, ideally, a single principle), which in turn cannot be derived from, or subsumed under, any higher principles. The “systematic in our knowledge,” Kant later states in the first Critique, consists in “its connection out of *one* principle” (A645/B673, emphasis added).

Against this, one might object that Kant cannot simply *assume* that reason seeks the absolute unity of knowledge, since this preoccupation with systematic unity already reflects a certain metaphysical picture. According to this broadly Platonic picture, there lies under the chaotic surface of empirical phenomena an eternal order that is to be discovered by reason alone. It would not be surprising that the conception of reason motivated by this metaphysical picture would necessarily lead to metaphysical thinking. Kant’s argument for the *necessity of metaphysics* thesis would turn out to be circular. It is thus important to note that the interest in absolute unity and highest principles can be seen as a consequence of reason’s concern with what I have called “inferential explanations.” If reason is to answer why-questions concerning universal propositions by locating them in the inferential net of our knowledge, it may be described as following an iterative procedure: If a proposition A is given, one must find premises from which A can be derived; if one has found suitable premises for A, one must find premises from which the premises of A can be derived, and so on. This process comes to a non-arbitrary hold only if one finds a premise which is so general that it cannot be derived from any further premise.<sup>9</sup> In this way, the search for inferential explanations automatically becomes a search for systematic unity, since reason completes its task only when it arrives at a highest principle, at which point the search for further “conditions” no longer makes sense. If giving inferential explanations is a task of reason, then it seems to follow that reason will have to look for ever more abstract and general principles until it finds the one single principle under which all our knowledge can be subsumed.

Up to this point all we have is what Kant calls a “logical maxim”—that is, a rule that directs reason in its logical employment and as such has no ontological import. It is only in the next sentence that Kant makes the shift from epistemology to metaphysics: “But this logical maxim [to look for the unconditioned] can only become a principle of *pure reason* through our assuming that if the conditioned is given, the whole series of conditions,

---

<sup>9</sup> It is not clear to me whether Kant wants to characterize the highest principle which stops the regress merely in terms of generality or also in epistemic terms (e.g. self-evidence).

subordinated to one another -- a series which is therefore itself unconditioned -- is likewise given, that is, is contained in the object and its connection" (A307/B364). Here the "logical maxim" is transformed into the metaphysical conditioned-unconditioned principle that no longer speaks of knowledge, but of the "object and its connection." This is the decisive move in Kant's argument for the necessity of metaphysics. It raises a number of difficult questions, only some of which I can address here.

The first difficulty concerns the precise content of the conditioned-unconditioned principle. In particular, what kinds of conditions is Kant talking about: causal or logical conditions? Necessary, sufficient, or necessary and sufficient conditions? Concerning the causal/logical distinction, it seems most plausible that Kant intends the term "condition" to cover both, since a little later Kant puts the conditioned-unconditioned principle as follows: "that the series of conditions (whether in the synthesis of appearances, or even in the thinking of things in general) extends to the unconditioned" (A308/B365). While the conditions in the "synthesis of experiences" presumably are causal conditions, or more generally, conditions of the *reality* of appearances, conditions in the "thinking of things in general" are mere logical conditions.

The next question is whether Kant thinks of necessary, sufficient, or necessary and sufficient conditions. Astonishingly, Kant nowhere is explicit about this. All three readings have some initial plausibility: When something is conditioned, it follows analytically that all of its necessary conditions are given. And from this one might be tempted to go on and conclude that either there must be a final necessary condition which itself is unconditioned or that the series of necessary conditions is infinite and, considered as a whole, is unconditioned. On the other hand, if something conditioned is given, it also seems to follow that there is *some* sufficient condition. And again, one might be tempted to go on and conclude that the series of sufficient conditions must end with something unconditioned or be infinite and thus itself unconditioned as a whole.<sup>10</sup> Finally, one might argue that Kant must have meant necessary and sufficient conditions, because on the one hand, Kant takes as his paradigm the structure of a syllogism, where the premises are logically *sufficient* conditions for the truth of the conclusion. On the other hand, Kant says that reason moves upwards from the conditioned to the condition, which move is logically valid only with respect to *necessary* conditions. So what Kant needs are conditions which are both necessary and sufficient. And although Kant's

---

<sup>10</sup> Read in this way, the conditioned-unconditioned principle would still not be the same as Leibniz' principle of sufficient reason which says of *everything* that it must have a sufficient reason. Kant's principle, read in terms of sufficient conditions, would only claim that if something is *conditioned*, it follows analytically that there must be some sufficient condition for it.

texts leave some room for disagreement here, I think that on balance they favour an interpretation in terms of both necessary and sufficient conditions. Kant's focus, in discussing the conditioned-unconditioned principle, does not lie on logical entailment, but rather on *knowledge* and hence on justification, rational explanation and reason-giving. In looking for the unconditioned, what reason is looking for is a kind of explanation for a given fact that does not leave any open questions. And this can be given only in terms of conditions which are both necessary and sufficient.<sup>11</sup>

A further question is whether Kant is correct to treat an infinite series of conditions, each of which is conditioned, as something unconditioned. Kant's idea seems to be that someone who claims, for instance, that the series of causes and effects extends infinitely into the past presupposes a *totality* of causes and effects just as much as someone who holds that there must have been a first uncaused event. Maybe it is misleading to say that an infinite series of causes and effects is *unconditioned*, since such a series may not seem to be the kind of object to which the distinction between conditioned and unconditioned can be applied. But what Kant means is that such an infinite series would have to be considered as a *totality* – as something that includes all past events as its members and thus cannot itself depend on any past event. So if we admit that a totality, as such, is unconditioned in this sense (cf. A322/B379), I think we can also accept Kant's claim that an infinite series of conditions must itself be unconditioned.<sup>12</sup>

More importantly, however, we must ask whether Kant has really shown that the conditioned-unconditioned principle flows naturally from the logical employment of reason. First, note that Kant does not claim that the move from the logical maxim to the metaphysical principle is necessary. Rather, he states that the logical maxim can become a principle of pure reason *only* if the conditioned-unconditioned principle is assumed. This would leave open the possibility that we retain the logical maxim as something like a regulative principle, but reject the metaphysical or constitutive principle as invalid; indeed, this is Kant's own suggestion in

---

<sup>11</sup> If, for example, we want an explanation for some event E, and we learn that some sufficient condition for E obtained, we still may wonder whether E would not have happened even if this condition had not obtained, since there might have been some other sufficient condition which would have brought about E anyway. If all we learn is that some necessary condition for E obtained, by contrast, we still do not know why E actually happened. Only if we know both necessary and sufficient conditions for E do we fully understand why E came about.

<sup>12</sup> In this way we can also deal with the problem, apparently unnoticed by Kant, that even in the finite case, there are two possibilities: Either there is at least one condition which does not depend on any further condition and hence is unconditioned, or the conditions form something like a circle of dependence, so that A is a condition of B, B of C and so on, until X is a condition of Y and Y of A. In the latter case, there would be a totality of conditions, but none of its members would be unconditioned. Still, the totality of conditions as a whole might be considered as unconditioned in the relevant sense.

the “Appendix to Transcendental Dialectic” (A648/B676).<sup>13</sup> But if Kant wants to maintain the *necessity of metaphysics* thesis, then reformulating the conditioned-unconditioned principle as merely regulative can only be a secondary move. He must first show that it is most natural (and, without a critique of pure reason, even unavoidable) to accept the principle in its constitutive form.<sup>14</sup>

Thus the question is whether someone who follows the logical maxim of looking for unconditioned principles of knowledge should be necessarily tempted to accept the constitutive principle that if the conditioned is given, then so is the unconditioned. I think that the answer must be “no.” The conditioned mentioned in the logical maxim is a piece of knowledge that reason seeks to subsume under more general knowledge; in the conditioned-unconditioned principle, by contrast, the unconditioned is meant to be “contained in the object and its connection.” The logical maxim speaks of conditions of *knowledge*, while the conditioned-unconditioned principle speaks of conditions of *objects* (and events). Even if we grant that looking for conditions of knowledge, or for general premises, takes us all the way to the idea of unconditioned knowledge, it is unclear why this should tempt us to posit the existence of some unconditioned *object* such as the world or God. The kind of metaphysical reasoning at work in the antinomies, for instance, nicely fits the pattern of the conditioned-unconditioned principle: If there is something conditioned, such as a point in time, a limited space, or an effect, it seems to follow that the complete series of its conditions (of earlier points in time, of neighbouring spaces, or of the effect’s causes) must exist as well. Here, the reasoning moves from individual objects and events to the *necessary* conditions of their existence. But it is difficult to see how this kind of reasoning about objects and their conditions might contribute to the project of unifying our knowledge by subsuming it under the highest principles, which project proceeds from general propositions to premises that are logically *sufficient* conditions of their truth.<sup>15</sup> The kinds of reasoning employed in the systematisation of knowledge on the one hand and in providing dialectical proofs of metaphysical theses on the other seem to be quite different. And because of this difference,

---

<sup>13</sup> However, Kant also says that the logical principle of the unity of reason presupposes a transcendental principle (A650/B678); on this see Michele Grier, *op. cit.*, Ch. 8.

<sup>14</sup> This points to a general tension in Kant’s views about metaphysics: on the one hand, the transcendental illusion that there must be something unconditioned is said to be unavoidable; on the other hand, explicit falsehoods and unwarranted metaphysical beliefs can be avoided by a critique of pure reason. Michelle Grier (*op.cit.*, Introduction) shows how this tension can be resolved if we insist (with Kant) that transcendental illusion, although they are unavoidable, need not deceive us (as long as we keep in mind the transcendental distinction between appearances and things in themselves).

<sup>15</sup> Maybe one could say that the theses and antitheses of the antinomies pretend to be such highest principles. But then they are not arrived at in the way Kant outlines in the Introduction to the Transcendental Dialectic, that is, they are not arrived at by moving via syllogisms from more particular to more general principles.

the logical maxim of systematizing our knowledge under as few principles as possible does not seem to motivate the acceptance of the conditioned-unconditioned principle.

However, this problem may be circumvented by making the plausible assumption that Kant, in the Introduction to the Transcendental Dialectic, intends the conditioned-unconditioned principle to apply to both: a) reasoning from conditioned knowledge to highest principles; and b) reasoning from conditioned objects to the unconditioned totality of their conditions. If we want to keep these issues separate, as I think we must, we will have to distinguish between *two* logical maxims and *two* corresponding metaphysical principles: One maxim directs reason to systematize knowledge by looking, with regard to each piece of knowledge, for a more general proposition under which it can be subsumed. Because following this maxim results in an iterative process (looking for ever more general propositions), it allows us to terminate this search only when all our knowledge has been subsumed under one ultimate proposition. In following this maxim, therefore, we must assume that a complete unity of knowledge under one principle can indeed be achieved. And from this epistemological assumption it is but a small step to the metaphysical claim that nature as the object of our knowledge is a totality, since only knowledge of a totality allows for complete systematisation. (The last step of the argument is not meant to be logically watertight, but merely plausible enough to explain why metaphysical thinking comes naturally to rational beings.)

The claim that nature as the object of our knowledge is a totality that allows for systematic representation is obviously not equivalent to the conditioned-unconditioned principle. In order to derive this latter principle, Kant would have had to start from quite another logical maxim—namely, that if a conditioned object or event is given, one must look for its condition, where the condition is not a general proposition but some other object or event. The problem with this maxim, from Kant's point of view, is that it does not address itself to reason as Kant defines it. Reason is the faculty of mediate inference or syllogism, and the relation between conditioned and conditioning objects and events is not the same as the relation between the premises and conclusion in a syllogism. To find relations of conditional dependence among individual objects and events is not the task of reason, according to Kant, but that of the understanding. So if the conditioned-unconditioned principle arises from ontologizing the maxim that directs us to look for the conditions of conditioned objects and events, then something must be wrong with Kant's claim that only reason, and not the understanding, gives rise to metaphysical illusions.

If we set this problem aside as concerning merely classificatory questions about our mental faculties, however, we can see how to explain the intuitive appeal of the conditioned-unconditioned principle. If we are to look for the conditions of conditioned objects and events, we must also look for the conditions of these conditions, whereupon we may find that the conditions themselves are conditioned. The maxim to find the condition for conditioned objects and events thus directs us to terminate our investigations only when we arrive at a condition that is unconditioned. If we follow this maxim, we must indeed assume that, for every object and event that is conditioned, there can be found a condition that is itself either conditioned or unconditioned. Again, from this assumption it is but a small step to the metaphysical principle that if something conditioned is given, so is the complete series of its conditions which itself is unconditioned.

Therefore, we have two quite different metaphysical principles that grow naturally out of rational thinking broadly conceived. One is the principle of the unity of nature that Kant discusses in the “Appendix to the Transcendental Dialectic,” a principle closely related to the principle of complete determination from which Kant derives the concept of God. The other is the conditioned-unconditioned principle, which is operative mainly (but not exclusively) in the antinomies. What both principles have in common, and what seems to me to be the real source of metaphysical thinking that Kant discovered, is that they are motivated by a certain *iterative* structure of reasoning. Thus in its attempt to discover inferential relations among given pieces of knowledge, reason follows an iterative procedure defined by the following rule: For any general proposition P known to be true, search your knowledge for some other general proposition from which P can be derived, until you arrive at a proposition from which all other knowledge can be derived. This procedure demands of us that we pursue our investigation indefinitely until we have completed the task of unifying knowledge under one principle. Another way to make the same point is this: If A and B are general propositions, and “Why A?” is good question to which “B” is the answer, then “Why B” is a good question, too, unless B is a proposition from which all the rest of our knowledge can be derived. This iterative structure of why-questions initiates a regress that is potentially infinite, unless this process can be concluded by the discovery of one ultimate principle of knowledge.

Similarly, the conditioned-unconditioned principle is motivated by reasoning that follows an iterative procedure: For any given object (event) that is conditioned, look for the object (event) which is its condition (e.g. its cause), until you arrive at an object/event which does not have a condition. Again, the same point can be made in terms of why-questions: If A and B are two events, and “Why did A happen?” is a good question to which

“Because of B” is the answer, then “Why did B happen?” is a good question, too, unless B is an event that is somehow self-explaining.

As Kant says in the A-Preface, the business of reason “must remain forever unfinished, *since the questions never cease*” -- unless reason makes recourse to transcendental assumptions (cf. A XIII; emphasis added). We now can see that ‘the questions never cease’ precisely because the iterative structure of reason allows questions to come to an end only at an ultimate principle, an unconditioned condition, an unexplained explainer. Metaphysical arguments have exploited this feature of rational thinking at least since Aristotle’s proof of a prime mover unmoved. It is the iterative structure of reasoning which “drives” reason from the conditioned to the unconditioned. This structure demands, in effect, that there be only two ways in which reason can be satisfied and thus only two ways for a theoretical claim to be rational: Either the claim is *ultimate*, in that it does not allow for further why-questions, or it can be derived from, or is otherwise suitably connected to, a claim that is ultimate. As I will argue next, Kant’s practical transformation of metaphysics transcends this conception of reason by disclosing a third way in which a theoretical claim can be rational.

## 2. The Practical Transformation of Metaphysics

Kant presents the central ideas of his practical transformation of metaphysics for the first time in the “Canon of Pure Reason” in the first Critique, but they come to full fruition only in the *Critique of Practical Reason* and the *Critique of Judgment*. Here I want to focus on Kant’s account of the postulates of pure practical reason, and on his thesis of the primacy of practical reason in the “Dialectic” of the second *Critique* (5: 107-48). Since I am interested primarily in the structural aspects of Kant’s argument and what they reveal concerning his conception of reason, I will restrict myself to a broad and somewhat schematic outline of Kant’s argument.

Pure practical reason, as Kant somewhat surprisingly points out in the second *Critique*, also leads us into dialectical fallacies. These fallacies arise from the same principle that lies at the core of the fallacies of pure speculative reason—namely, the conditioned-unconditioned principle: “As pure practical reason it [reason] likewise seeks the unconditioned for the practically conditioned[...]; and this unconditioned is [...] sought as the unconditioned totality of the object of pure practical reason, under the name of the *highest good*” (5: 108). As Kant explains, the term “highest good” is ambivalent, since it can either mean the unconditioned condition of all that is conditionally good (*bonum supremum*), which is the moral law, or the complete good (*bonum consumatum* or *perfectissimum*), which is the conjunction of perfect

morality with perfect happiness (5: 110). It is the highest good in the latter sense that interests Kant. Now the notion of a highest good gives rise to an antinomy which, deviating slightly from Kant's own presentation (5: 114), can be stated thus: It is both *necessary* and *impossible* for us to bring about the highest good: It is practically *necessary* (i.e. obligatory), because as finite rational beings we must strive for both moral perfection and happiness. But the realization of the highest good is also *impossible*, because both conceptually and empirically moral perfection and happiness just don't go together.

Kant gives no extended argument for his claim that bringing about the highest good is *necessary*. All he says is that, from both the partisan perspective of the person concerned and from the perspective of a disinterested reason, morality without happiness is insufficient as an end of our will. But this alone does not mean that we must be able to *realize* this end. Perhaps an argument to this effect can be based on Kant's claim in the *Groundwork* that happiness is an end finite rational beings must have "by natural necessity" (4: 415). If we add to this that to strive for moral perfection is a categorical imperative, it follows that finite rational beings must strive both for moral perfection and for their own happiness, which striving makes sense only if we believe that its end, the highest good, can at least in principle be realized. But still this argument is unconvincing. It may be necessary that I *wish* to achieve it; but if I don't believe I can reach it, it follows that I can't even try. So there's nothing irrational in giving up the aim of perfect happiness; quite to the contrary, for a finite being it would be highly irrational to try to reach perfect happiness, since this endeavour is bound to fail. Rather, it seems reasonable to consider perfect happiness (as well as the highest good of which it is a part) as something like regulative ideas of pure practical reason – necessary objects of our will that regulate our actions even though they can never be fully realized. But then it does not seem to be necessary to assume that we can *realize* the highest good.

If we accept that claim for the sake of the argument, however, we must turn to the other side of the antinomy, namely that it is *impossible* to realize the highest good. Here Kant reasons as follows: In order to realize the highest good, I must either bring about moral perfection by striving for happiness, which is, as Kant says, "absolutely impossible," or I must bring about my happiness by striving for morality, which is empirically impossible (5: 113). Kant does not consider the possibility that one strives for both morality and happiness independently of each other. His reason is that both are supposed to be elements of the highest good and hence must be related either logically or causally. Since they are not related logically (one doesn't follow from the other), they must be related causally (one bringing



about the other). If we accept this, it follows that the highest good cannot be realized. Hence, the antinomy.

Kant resolves this antinomy by disputing the assumption that existence in the world of sense is “the only mode of existence of a rational being” (5: 114). In the absence of this assumption, Kant claims, it turns out to be possible to realize the highest good after all. Although it may be empirically impossible to realize it directly, it is possible for rational beings, as members of an intelligible world, to achieve the highest good indirectly, in that their moral disposition motivates an intelligible author of nature to bring about their happiness. The idea is that if we do our part in realising the highest good by leading morally perfect lives, it is at least possible that God will do his part and award us perfect happiness. With this conclusion in hand, Kant goes on to derive our immortality and the existence of God as “postulates of pure practical reason.” According to Kant, we cannot do our part directly, simply by leading a morally perfect life, because for the sensible beings we are moral perfection is impossible. Since the moral law nevertheless categorically demands that we strive for moral perfection, we must assume that we can at least approximate it in an infinite process of moral improvement, the infinity of which requires our immortality. God, on the other hand, can do his part in awarding the worthy with happiness only if he exists in the first place. So the practical necessity of the highest good forces us to postulate our own immortality and the existence of God.

This summary of Kant’s argument is meant only as a brief reminder. Several steps of the argument are problematic and would deserve closer attention than they can be given here. What interests me here are not the details of Kant’s argument and whether or not it is convincing, but rather what kind of argument Kant presents and the epistemic status he assigns to its conclusion. Interestingly, Kant himself interrupts the argument in several places with epistemological and meta-philosophical reflections, the most important of which, at least for our purposes, is the section on the “Primacy of Pure Practical Reason in its Association with Speculative Reason” (5: 119-21). The primacy of one faculty over another, Kant explains, consists in the fact that the “interest” of one faculty is subordinated to the interest of another. The interest of pure speculative reason, Kant goes on, is “knowledge of objects up to its highest a priori principles” (5: 119f.), which corresponds precisely with what Kant says in the *Critique of Pure Reason*. The interest of pure *practical* reason, by contrast, is the “determination of the will with respect to the final and perfect end,” that is, with respect to the highest good (5: 120). Now the question is which interest—that of speculative reason or of practical reason—has primacy. Kant explicitly assumes that the interests of speculative and

practical reason cannot conflict and that their results cannot contradict each other, since “this [viz., consistency] is a condition for having reason at all” (ibid.) Speculative reason has primacy if “practical reason may not assume and think as given anything further than what speculative reason affords from its own insight” (ibid.). By contrast, practical reason has primacy if speculative reason must accept theoretical propositions which it can on its own neither prove nor disprove, but which are necessary presuppositions of a principle of pure practical reason:

“But if pure reason of itself can be, and really is, practical, as the consciousness of the moral law shows it to be, it is only one and the same reason which judges a priori by principles, whether for theoretical or practical purposes. Then it is clear that, if its capacity in the former is not sufficient to establish certain propositions positively (propositions which however do not contradict it), it must accept these propositions just as soon as they belong *inseparably to the practical interest* of pure reason. It must accept them indeed as something offered from the outside and not grown on its own soil, but still as sufficiently warranted [...]” (5: 121).

So Kant concludes that pure practical reason has primacy over pure speculative reason (ibid.), since, as Kant states at the end of the paragraph, “every interest is ultimately practical, even that of speculative reason being only conditional and reaching completion only in practical use” (ibid.).

What Kant has argued for in the above-cited passage is merely conditional: If there are theoretical claims that are necessary presuppositions of principles of pure practical reason, then speculative reason must accept them “as sufficiently warranted.” Now the postulates of immortality and of the existence of God fit precisely the demands of the primacy of practical reason, since a postulate of pure practical reason, on Kant’s definition, is a “theoretical proposition which is not as such demonstrable, but which is an inseparable corollary of an a priori unconditionally valid practical law” (5:122). According to Kant, then, we are rationally justified in believing in our own immortality and in God’s existence as *theoretical* truths, even though we do not *know* these truths, because they are necessary implications of considering ourselves bound by the moral law. Put more generally, according to Kant’s thesis of the primacy of practical reason, theoretical claims (as long as they do not conflict with known theoretical truths) can be “sufficiently warranted” by being implied by norms we consider as

absolutely binding. Hence it is rational to accept them, even if their truth cannot be established theoretically (that is, even if they do not constitute theoretical knowledge).

### 3. Conclusion

If we now look back at Kant's genealogy of metaphysics in the first Critique, we can see that the kind of rational warrant Kant acknowledges with the primacy of practical reason does not fit the description of reason as governed by iterative procedural maxims. In the Introduction to the Transcendental Dialectic, Kant traces the necessity of metaphysics to the iterative structure of rational thinking that requires us to look for the condition of the conditioned, and for the condition of the condition, and so on, until we find something unconditioned. As we have seen, the details of Kant's derivation of the conditioned-unconditioned principle raise numerous exegetical and philosophical questions. Kant, perhaps for expository reasons, did not distinguish properly between conditions that may be traced to the highest principles or first foundations of *knowledge*, and those conditions that lead us to some unconditioned *object* or event. But I think the general thrust of Kant's argument is both clear and convincing: That is, there is a tendency in rational thought not just to ask questions, but to react to each answer with a further question: "Why A?"- "Because of B." – "But why B?" and so on. In the context of philosophy, the tendency to not be content with answers that allow for further questions can seem absolutely compelling. It makes us look for ever more general propositions and prior conditions until we reach a supreme principle or an unconditioned condition. The effect is that a proposition is rationally acceptable only if it is either the highest principle itself, or if it can be somehow derived from such a principle.

What Kant has recognized in the *Critique of Pure Reason* is that this tendency is at the root of (at least one important strand of) metaphysical thinking, since once we give way to it, we will either get lost in the infinity of ever higher conditions, or we arrive at beliefs in something unconditioned. But it seems that Kant was wrong to assume that this kind of thinking is peculiar to human reason as such, as his own thesis of the primacy of practical reason allows us to see. As Kant himself acknowledges in the *Critique of Practical Reason*, there is a third way in which a proposition can be acceptable to pure reason--namely, by being indispensable for certain practical purposes.

Kant is, I think, absolutely right here. It would be irrational not to accept the necessary preconditions of those things which matter most to us – of our values, but also and more generally of our ability to think of ourselves as morally responsible agents. We may disagree with Kant concerning what these presuppositions are; while I agree that freedom of the will belongs in the list of postulates, Kant’s arguments for the postulates of God and immortality do not convince me. What matters here, however, is that Kant presents a *kind* of argument, based on the primacy of practical reason, which goes beyond the traditional conception of reason haunted by the alternative of infinite regresses or unconditioned conditions. What halts the regress of conditions is not necessarily something unconditioned, but rather the *practical* necessity of a certain belief.<sup>16</sup>

Of course, people have always held fast to the truths upon which their practical lives depend. But philosophers have traditionally disparaged this kind of trust as irrational gullibility that may suffice for simple folk, but not for philosophy. Kant breaks with this tradition. While he of course insists that not every kind of wishful thinking is rational (cf. 5: 143-4 fn.), he argues that where presuppositions of our conception of ourselves as moral agents are concerned, reason requires us to believe in their truth even without theoretical proof (as long they do not contradict theoretical knowledge). It is important to see that this kind of practically validated belief, according to Kant, is not in any way second rate. Since the moral law is “in itself absolutely certain” (5: 142), it does not need any underpinning from theoretical reason -- and neither do the postulates based on it. Their rational warrant is different from, but not inferior to, empirical truths and transcendental conditions of the possibility of experience. This seem to me to be the revolutionary import of Kant’s doctrine of the postulates: It is just as rational to believe in the necessary preconditions of our practical identity as it is to believe in theoretically established truths.

Kant was perhaps the first philosopher to have developed a clear alternative to the traditional philosophical conception of reason. This alternative is based on the idea that a *theoretical* belief can be rationally warranted by being *practically* indispensable. Kant limited this kind of rational warrant to what is indispensable for considering oneself as bound by the moral law. But I think we may generalize this insight and insist that it is rational to believe in the presuppositions of *all* the things that really matter to us, as long as the resulting belief

---

<sup>16</sup>To be sure, the objects of the postulates of pure practical reason are, precisely, Kant’s prime examples of something unconditioned, namely God, freedom, and an immortal soul. My point is, however, that belief in these objects is not based on climbing up the ladder of conditioned objects and their conditions until we reach something unconditioned, but rather on the practical indispensability of the belief in question. This is a “shortcut” to the unconditioned not envisaged in Kant’s conception of reason in the “Introduction to the Transcendental Dialectic” of the first *Critique*.

does not contradict empirical evidence. This is a consequence accepted by philosophers as diverse as Dewey, Heidegger, and Wittgenstein – philosophers not generally considered as Kantians. But if I am right, it was Kant who led their way.