The Unity of Reason: Kant’s Copernican Presupposition

Dr Edward Thornton

In the controversial Appendix to the *Transcendental Dialectic*, Kant claims to “complete the critical work of pure reason” [A670/B698] by providing a transcendental deduction of the ideas of pure reason. In order to analyse the role that this Appendix plays in the first Critique, this paper will read the Appendix alongside Kant’s comments in the B-Preface concerning the astronomy of Copernicus. Through an analysis of the nature of Kant and Copernicus’ respective use of presuppositions, and by looking at their respective attempts to unify a science around a single systematic conception of the object of that science, this paper will offer a defence of the fundamental role that Kant’s transcendental deduction of the ideas of pure reason plays in the text.

Keywords: Kant, Copernicus, Reason, Presupposition, Metaphysics

1. Introducing the Transcendental Presupposition of the Unity of Reason

During Kant’s critical turn, something surprising happens to the nature of philosophical presuppositions. It had been common practice in Western philosophy, at least since Descartes, to avoid presuppositions at all costs; nothing should be taken for granted in argument that cannot be defended on its own grounds, lest it turn out to be untrue and jeopardise the whole endeavour. Kant’s first Critique initially repeats this disciplinary norm by warning the reader that any form of presupposition – be it hypothetical or axiomatic – should be “treated as contraband” [Axv]. To put this in the terminology of Kant’s first Critique, we might say that, to provide a foundation for metaphysics we must offer an analysis of what can be known *a priori* concerning the necessary conditions of experience and that it is only by avoiding any reliance on contingencies that we can reach our conclusions with apodictic certainty. However, unlike the vast majority of thinkers before him, Kant regularly found himself providing
arguments in which presuppositions come at the end, rather than the beginning, of an argument, filling the place where we would normally expect to find the conclusion. To take the *Critique of Pure Reason* as a whole, we can say that Kant’s aim is to show that, starting from what is given in experience, it is possible to determine the necessary preconditions for the possibility of such experience. Another way of putting this is to say that, if we begin with an analysis of the structure of experience, then it is possible for us to determine what *must be presupposed* in order to account for the possibility of such a structure.¹ In this sense, Kant distinguishes between two different kinds of presuppositions: there are those which come at the beginning of an argument and serve as an undefended and therefore dangerous ground for argumentation, and there are those which come at the end of an argument and tell us what must be presupposed in order to account for something that is given.² Let us call the former of these ‘prefatory-presuppositions’ and the latter ‘conclusory-presuppositions’. In the *Critique of Pure Reason*, Kant regularly finds himself attempting to unearth and dispel prefatory-presuppositions and to replace them – often via the use of transcendental arguments – with conclusory-presuppositions.³

This format recurs throughout the first Critique in relation to the three fundamental faculties of sensibility, understanding, and reason. In regard to sensibility, Kant shows that while we cannot presume that things-in-themselves have any specific form, we can recognise that all sensibility is structured by the forms of space and time. Because space and time are the essential forms of

1 Howard Caygill captures this in his formulation of Kant’s attempt to ground the possibility of synthetic *a priori* judgements: “[T]he conditions of the possibility of synthetic *a priori* judgements comprise principles which unite conceptual and intuitive elements, and which are presupposed by discrete acts of synthetic *a priori* judgement” (Caygill, 2009, 385).
2 Throughout this section, and the rest of the paper, the term presupposition is used in a purposefully broad sense. For example, my usage of the term ‘presupposition’ is broader than that used by Strawson (1952, 175-179), because I am concerned not only with logical presuppositions which appear as supressed premises in individual arguments but also with metaphysical presuppositions, which underlie belief systems rather than arguments.
3 I have avoided naming the latter of these ‘transcendental presuppositions’ because, as we shall see, Kant uses this terminology in a very specific way, which has a narrower sense than my ‘conclusory-presuppositions’.
our sensibility, we can be sure that all objects presented to us in experience – that is to say, all ‘appearances’ – will be both spatially and temporally organised. Thus, at the end of the *Transcendental Aesthetic*, we can conclude that space and time are “pure a priori intuitions”, which we are justified in presupposing as the conditions for all possible objects of experience [B73]. Here we have replaced the prefatory presupposition that every object is spatially and temporally organised, with the conclusory presupposition that every object of experience is spatially and temporally organised. Either of these presuppositions could act as a ground for our scientific examination of the nature of reality, but while the former is ungrounded, the latter can be given a ground. Turning to the faculty of the understanding, in the *Transcendental Logic*, Kant argues that as well as there being essential forms of the receptivity of impressions, there are also certain regular features of our conceptual grasping of experience. Just as possible objects of experience must conform to our faculty of sensibility, possible objects of knowledge must conform to the categories of our understanding. Specifically, Kant argues that we can specify twelve “pure concepts of the understanding which apply a priori to objects of intuition in general”, organised under the four categories of “quantity”, “quality”, “relation”, and “modality” [A79-80/B105-106]. It is not my intention here to offer a defence of Kant’s deduction of these categories, but simply to point out that the methodology used by Kant to provide such a deduction comes to a resolution with conclusory-presuppositions.

The third case, regarding the faculty of reason, is more complex and it is here that I would like to focus our attention. In Kant’s analysis of this faculty, he follows the same argumentative pattern as in the case of sensibility and understanding. In the final sections of the *Transcendental Dialectic*, Kant sets out to show that reason produces its own illusions when it steps beyond the bounds of its rightful domain. Kant then tests the exact boundaries of this domain via the use of the paralogisms, the antinomies, and the analysis of the Ideal. Here he
shows that reason cannot legitimately apply itself to anything which falls outside of the bounds of possible experience. As such, there can be no proof in psychology of the existence of the unified thinking subject, no proof in cosmology of the unity of the world as a whole, and no proof in theology of the existence of God. Despite the impossibility of such proofs, Kant nonetheless argues that these three ideas – namely the self, the world, and God – must be presupposed as regulative ideas in order to allow reason to function. Kant refers to these as “transcendental ideas” [A311/B368] or “ideas of pure reason” [A669/B697].

To make things slightly more complex, Kant also discusses the need for a more general regulative idea of pure reason, which conditions the three ideas introduced above. Kant is clear throughout the *Critique of Pure Reason*, that the faculties of sensibility, understanding, and reason all support one another in various ways. In the case of the latter, Kant will argue that reason serves to unify the system of concepts which the understanding must use to unify the object of experience: “Just as the understanding unifies the manifold in the object by means of concepts, so reason unifies the manifold of concepts by means of ideas” [A644/B672]. Thus, the three transcendental ideas of pure reason discussed above are presupposed specifically to allow for the synthesis of the concepts of the understanding. Kant adds to this the claim that reason carries out this task by “positing a certain collective unity as the goal of the activities of the understanding” [A644/B672]. For Kant, the three regulative ideas of the subject, the world, and God, can only operate as regulative ideals at all if reason seeks to unify the understanding. As Kant continues:

“The law of reason which requires us to seek for this unity, is a necessary law, since without it we should have no reason at all, and without reason no coherent employment
of the understanding, and in the absence of this no sufficient criterion of empirical truth” [A651/B679].

Here Kant has provided a kind of transcendental argument for the existence of a law that requires pure reason to at least attempt to unify the concepts of the understanding. In brief, Kant is claiming that the understanding could not systematically organise all possible objects of knowledge under a single set of necessary categories, unless reason put the understanding to work in a specific way. The three ideas of pure reason mentioned above can only act to unify the concepts of the understanding if reason uses them in an attempt to seek for this unity. It is for this reason that Kant will write: “In order, therefore, to secure an empirical criterion we have no option save to presuppose the systematic unity of nature as objectively valid and necessary” [A651/B679]. Effectively, in this section of the Critique, Kant posits the necessity of a regulative ideal even more general than the three transcendental ideas of the unity of the subject, of the world, and of God. To these, Kant has added the requirement that we presuppose the unity of the whole of knowledge and the concomitant presupposition of “the systematic unity of nature”. Kant speaks of this presupposition as a “transcendental presupposition” [A651/B679] and, as I will attempt to show, it has a very special character, which differentiates it from the other presuppositions that Kant makes in the book.

---

4 The precise way in which these two presuppositions are concomitant will not be completely clear until after the analogy with Copernicus has been completed in section 4. However, my argument here proceeds by showing that it is not possible to presuppose the unity of knowledge without also presupposing the unity of nature. It may also be worth noting here that there is some evidence that Kant changes his view on this matter in the third Critique. According to Paul Guyer, in the first Critique Kant argues that “reason is allowed to add the ideal of systematicity to the theoretical cognition produced by the understanding”, because in order to seek systemicity in our concepts of nature “we have to assume that nature itself is systematic” (2003, 281-282); but that in the third Critique Kant argues that “the necessary accordance of ‘experience in general’ with the ‘transcendental laws of the understanding,’ which is enough to assure us that some concept and law or other must in principle exist for any empirical intuition we encounter, is not itself enough to assure us that we can actually find a law for any empirical intuition” (2003, 284). In this paper I have left Kant’s later position on this matter to one side in order to concentrate on the arguments made in the first Critique.
2. An Analysis of the Transcendental Presupposition of the Unity of Reason

Unsurprisingly, this “transcendental presupposition” of the unity of reason, which connects the unity of knowledge and the systematic unity of nature, is a very contentious one. The controversy circles around two questions that we can ask of this presupposition. The first is whether – according to the terminology I have defined – we should consider this presupposition to be prefatory or conclusory. Is it the case that we must simply presuppose the unity of knowledge and the systematic unity of nature as a starting point for any possible metaphysics, or is it the case that we can provide a convincing argument for the necessity of taking up this presupposition? The second question concerns what Kant means when he says that “we have no option save to presuppose the systematic unity of nature” [A651/B679, my italics]. Is Kant making a direct claim about the actual systematicity and unity of nature, or a pragmatic claim about the need to assume the unity of nature in order to think? Is this “law” of the unity of reason supposed to apply to reality, or is it a kind of heuristic device that only determines the operation of reason?

Finding out whether Kant’s transcendental presupposition of the unity of reason is prefatory or conclusory should be a simple task: we simply need to check in the critique to see if Kant provides an argument that is intended to compel us to make this presupposition. In practice, however, things are not this simple. In the first section of the Appendix to the Transcendental Dialectic, Kant seems to rule out the possibility of providing a deduction for transcendental ideas. After speaking of the usefulness of the ideas of pure reason, including the idea of the unity of nature, Kant writes: “A transcendental deduction of them cannot however, be effected; in the case of ideas, as we have shown above, such a deduction is never possible” [A663/B691]. With this comment, Kant suggests that the transcendental presupposition of the unity of reason
is prefatory. It acts to direct the unifying activity of reason, which is required for the systematic operation of the understanding, but there is no way we can defend it on its own grounds. Shortly after this worrying admission, however, Kant seems to volte-face, recognising that without such a deduction of the ideas, they become nothing but unguarded hypotheses. Spurred on by the danger this would cause, Kant decides that it is necessary to provide a deduction of the ideas of reason after all, even if it must be of a different kind than that given for the categories of the understanding:

“We cannot employ an a priori concept with any certainty without having first given a transcendental deduction of it. The ideas of pure reason do not, indeed, admit of the kind of deduction that is possible in the case of the categories. But if they are to have the least objective validity, no matter how indeterminate that validity may be, and are not to be mere empty thought-entities (*entia rationis ratiocinantis*), a deduction of them must be possible, however greatly (as we admit) it may differ from that which we have been able to give of the categories.” [A669-670/B697-698]

The task of supplying a deduction of the ideas of pure reason, including the idea of the unity of reason itself, is not one that Kant takes lightly. He recognises that, the process of providing a transcendental deduction of space and time, as the forms of intuition [A87/B119-20], and of providing a transcendental deduction of the categories [A92/B124] and the concepts [A95, B130] of the understanding, would have been pointless if the activities of reason could not be given a similar ground. The reason for this is that, while subjective experience is possible in a limited sense without the activity of reason, the study of experience through the use of the empirical sciences is not possible without the systematicity provided by reason. It is useful to specify here what I mean by ‘experience’ when I say that without the application of reason
‘experience’ would not be possible. Gerd Buchdahl notes that for Kant the word ‘experience’ can be used in one of two ways: “‘Experience’ may refer to experience of single particular states of affairs, but also to that ‘systematic experience’, denoted by what is called ‘the theoretical corpus of knowledge’” (1967, 213). To say that ‘experience’ is not possible without the activity of reason is to use the concept of ‘experience’ in the latter way. For this reason, Kant will suggest that the task of providing a transcendental deduction of the ideas of reason is the final undertaking of the critique, which “will complete the critical work of pure reason” [A670/B698].

In the second section of the Appendix, Kant does provide a deduction of sorts for the ideas of pure reason. This very fact is enough to show that Kant takes these presuppositions to be conclusory rather than prefatory. The nature of these deductions, however, are somewhat at odds with the transcendental deductions given earlier in the Critique. In relation to sensibility and to the understanding, Kant had shown that the forms of intuition and the categories can be taken as a priori and necessary laws because they are involved in constituting the objects to which they apply. We can be sure that all objects of experience will be spatially and temporally ordered, because sensibility constitutes all possible objects of experience under the forms of intuition. We can be sure that all objects of possible knowledge will conform to the categories of the understanding, precisely because the understanding constitutes these objects according to the categories. The same cannot be said for the ideas of pure reason. Kant is clear that, while it is possible to provide arguments for the necessity of the self, of the world, and of God, these arguments turn on the practical necessity of such presuppositions, and not on their constitutive activity.⁵

⁵ See, for example, Kant’s claim that, “[w]e may not say that this idea is a concept of the object, but only of the thoroughgoing unity of such concepts, in so far as that unity serves as a rule for the understanding” [A645/B673]. Here Kant is pointing to the fact that the ideas of pure reason do not refer to objects of thought, but only play the
Before quoting Kant’s comments at length, it is important to clarify the manner in which Kant suggests we relate to these ideas. For Kant, the ideas of pure reason are not to be taken as objects of knowledge, but merely as a kind of direction for the activity of the understanding. When we refer to the self, for example, we do not relate to the self as an object which it is possible for us to know, but only as a “schema to which no object, not even a hypothetical one, is directly given” [A670/B698]. Here, the idea of the self only acts as a reference point, which will allow for the systematisation of the true objects of the understanding. Speaking of the use of the ideas of pure reason in this way, Kant writes that they serve “to represent to ourselves other objects in an indirect manner, namely in their systematic unity, by means of their relation to this idea” [A670/B698]. Kant’s use of the ideas of pure reason operates as follows: we must presuppose certain ideas in order to allow the understanding to operate in a systematic way; we should not take these ideas to refer to any object, but should accept them as regulative principles, which allow us to complete the project of reason. This is the account that Kant gives of the transcendental deduction of these ideas of pure reason:

“If, then, it can be shown that the three transcendental ideas (the psychological, the cosmological, and the theological), although they do not directly relate to, or determine, any object corresponding to them, none the less, as rules of the empirical employment of reason, lead us to systematic unity, under the presupposition of such an object in the idea and that they thus contribute to the extension of empirical knowledge, without ever being in a position to run counter to it, we may conclude that it is a necessary maxim of reason to proceed always in accordance with such ideas. This, indeed, is the transcendental regulative role of allowing the objects of cognition to be unified. To say that these arguments turn on the practical necessity of such presuppositions is to say that they are required for the practical use of reason.
deduction of all ideas of speculative reason, not as constitutive principles for the extension of our knowledge to more objects than experience can give, but as regulative principles of the systematic unity of the manifold of empirical knowledge in general.” [A671/B699]

This section of the Appendix is particularly useful for clarifying the relationship between the three transcendental ideas of pure reason – the self (psychological), the world (cosmological), and God (theological) – and the transcendental presupposition of the unity of reason. Kant claims that the three transcendental ideas are required because they lead us to “systematic unity”, but that this is only the case on the further “presupposition of such an object in the idea”. That is to say that further to the presupposition of the three specific transcendental ideas, we must also presuppose that the concepts of the understanding form a unity which allow our application of them to be systematic. As Kant states: “It is, indeed, difficult to understand how there can be a logical principle by which reason prescribes the unity of rules, unless we also presuppose a transcendental principle whereby such a systematic unity is a priori assumed to be necessarily inherent in the objects” [A651/B679]. The unity of reason is thus a further condition for the possibility of the systematicity of the understanding, on top of the application of the three more specific ideas of pure reason.

It is important to note that the “objects” mentioned in this quotation are not the objects of empirical knowledge, but the objects of the “rules” in question. Kant is clear that while the objects of the faculty of the understanding are the objects of knowledge, the object of the faculty of reason is the understanding itself: “Reason is never in immediate relation to an object, but only to the understanding… Reason has, therefore, as its sole object, the understanding and its effective application” [A644/B672]. When Kant says that in order for

---

6 For more on the relationship between these two ideas, see Zuckert (2017, 98-106).
reason to prescribe a unity of rules, we must presuppose a systematic unity of its object, the object in question is the understanding.

Kant has presented the transcendental ideas as merely regulative principles that are required to give reason its necessary direction, but simultaneously Kant has claimed that the presupposition that we are required to make of the unity of nature should be thought of as “necessarily inherent in the objects” [B679, A651]. As Peter McLaughlin points out, Kant seems to have presented us with “a principle that is regulative, not constitutive, but somehow claims to be transcendental, objectively valid and necessary” (McLaughlin, 2014, 561). This is a somewhat baffling claim. How can a regulative principle have objective validity? To make sense of this situation, we must remember that, for Kant, the central activity of the faculty of reason is a normative one. Reason’s task is not to provide true knowledge, but to provide justifications for the acquisition of such knowledge. As such, the idea of the unity of reason is not posited factually, but only as an aim. Reason does not so much take the concepts of the understanding to be unified, as it takes them to be unifiable. At this point at least, the transcendental presupposition of the unity of reason is presented not as a claim about the nature of the reality that is grasped by concepts, but as a normative instruction given by reason for the application of the understanding. As McLaughlin highlights: “Reason is the realm of the normative, not the factual, and the objects of the ideas that reason produces or collects are not things or facts but ideals or norms that the understanding is supposed to pursue” (McLaughlin, 2014, 559).

Keeping in mind this normative quality of the faculty of reason is helpful in dismissing what Lea Ypi terms the “weak reading” of Kant’s transcendental ideas. This weak interpretation makes the deduction of transcendental ideas “no big deal” because it claims that “the idea of
systematicity of nature only has a heuristic or methodological status, with no bearing on the way in which the categories of the understanding are applied to create empirical concepts” (Ypi, 2017, 167). This is a position implicitly accepted by Ralph Walker, among others, who claims that the deduction of the ideas of pure reason “does not seek to show that the world actually contains these kinds of unity and completeness” but only that “we are justified in proceeding as if it did” (Walker, 2006, 247). However, this view goes against the claims that Kant makes in the Appendix. As we have seen, Kant is clear that, while the idea of the unity of reason does not relate directly to any factual claim about the unity of nature, it is the case that we must presuppose a priori that the unity which reason seeks is “necessarily inherent in the objects” [A651/B679]. So, while the unity of reason is first introduced as a mere regulative principle, it is not possible to apply this regulative principle unless we actually take it to be the case that our concepts are unified, and thus that our knowledge may be systematic.

A close reading of Kant’s deduction of the transcendental ideas in the Appendix to the Transcendental Dialectic shows that we cannot defend a weak interpretation of the unity of reason. The approach left open to us – Ypi’s “strong reading” – maintains that “upon further scrutiny” Kant recognised that “reason’s heuristic/subjective use had to be grounded on a transcendental principle presupposing the conformity to ends of nature itself” (Ypi, 2017, 167). This is a position that I will be defending in the following sections of this paper. For now, however, all I want to do is offer an outline of this position by paying close attention to what it would mean to say that the transcendental idea of the unity of reason is constitutive in some sense.7 While the idea of the unity of reason is not constitutive of the objects of knowledge, we

---

7 Michelle Grier has written in detail about the role that the transcendental presupposition of the unity of reason plays in Kant’s first Critique. See especially ‘The Demand for Systematic Unity’ and ‘The Unity of Reason’ (Grier, 2004, 268-287). I agree with Grier that there is some ambiguity in Kant’s presentation, and also that a mere heuristic account of the demand for systematic unity is ultimately insufficient. In this sense my conclusions are also aligned with those of Sasha Mudd (2013). My approach to the problem differs from that of Grier on the grounds that I aim to show how it is ultimately Kant’s position as a metaphysician which makes the apparently
can still interpret it as constitutive of our system of knowledge. As we have already seen, Kant’s transcendental deductions of the forms of intuition and the categories of the understanding are constitutive in the sense that they constitute the objects to which sensibility and understanding are applied. We have also seen that Kant argues that the transcendental deduction of the ideas of reason should be understood “not as constitutive principles for the extension of our knowledge to more objects than experience can give” but as “regulative principles of the systematic unity of the manifold of empirical knowledge in general” [A671/B699]. What is telling here, however, is that Kant does not compare the roles of being constitutive and regulative per se. Instead, Kant only dismisses the claim that transcendental ideas are constitutive of the objects of experience. He does not dismiss the idea that they are constitutive of the object of reason. If the object of reason is the systematicity of the concepts of the understanding, then to say that the transcendental idea of the unity of reason is constitutive is to say that this idea is involved in the constitution, rather than discovery, of the systematicity of knowledge.8

The law of the unity of reason requires us to presuppose the possible unity of all knowledge and therefore the unity of nature. As Kant claims, this is “a necessary law” because “without it we should have no reason at all, and without reason no coherent employment of the understanding, and in the absence of this no sufficient criterion of empirical truth” [A651/B679]. Thus, while the transcendental idea of the unity of reason is not involved in the constitution of the possible objects of knowledge, it is involved in the constitution of the object of reason, namely the systematicity of the understanding. Kant’s use of the ideas of pure reason

8 My point here is that the systematic unity of knowledge does not logically precede reason’s unification of it. Knowledge can only be systematic to the degree that it is constituted by the activities of a reason which seeks unity in its object.
is therefore not merely heuristic. As McLaughlin points out: “While these principles are sometimes called ‘heuristic’ by Kant, they are not methodological suggestions justified by utility or success in practice. They are normatively constitutive of the rationality of scientific practice itself” (McLaughlin, 2014, 561, my italics). In summary, the transcendental deduction of the ideas of pure reason, including the idea of the unity of reason itself, should be understood as regulative and normative principles, which are not involved in the constitution of objects, but are nevertheless constitutive in the sense that they are necessary for the production of a systematic unity of concepts in the understanding.

On this reading, Kant’s transcendental presupposition of the unity of reason is conclusory rather than prefatory because it comes as the conclusion of a transcendental deduction, but it is unlike the others presented in Kant’s critique – including those which come as the conclusions to other transcendental deductions – because it relies on a set of normative claims. As such, the unity of reason should be understood as something which we are justified in presupposing, and which we must presuppose, if we are to think rationally. In order to provide more textual evidence for this interpretation, I will now turn to Kant’s infamous comments in the B-Preface, where he draws an analogy between his own project in metaphysics and that of Copernicus in astronomy. What I will show here, in direct contrast with the majority of interpretations of Kant’s analogy, is that the comparison with Copernicus is intended to shed light on Kant’s use of presuppositions and their role in the grounding of a science. Specifically, I will show that an analysis of this analogy can help to explain why the unity of reason must be presupposed as a

---

9 I say ‘merely’ heuristic because it is possible to conceive of a heuristic device that is also constitutive of a practice. This possibility is suggested by Walker’s analysis of transcendental arguments, in which he shows how arguments which prove the truth of a proposition $p$, and arguments which prove only that we rationally ought to believe that $p$, are ultimately equivalent (Walker, 2006, 260-261).

10 There is an emerging consensus in the literature that Kant’s transcendental deduction of the ideas of pure reason in the Appendix to the Transcendental Dialectic is a failure and that Kant returns to this task in the introduction to the Critique of Judgement. See Horstmann (1989), but also Ypi, (2017, 170) and McLaughlin (2014, 554). My analysis of the preface of the first Critique is an attempt to show that Kant’s initial transcendental deduction of the ideas of pure reason is more successful than it first appears.
regulative idea, without reducing it to a mere ‘heuristic’ principle. Furthermore, while I have already suggested that the presupposition of the unity of reason must be understood as including both the presupposition of the unity of nature and the presupposition of the unity of knowledge, by offering a defence of the non-heuristic role of Kant’s transcendental presupposition of the unity of reason we will see why the unity of nature and the unity of reason go hand-in-hand.

3. Copernicus and Kant on the Presupposition of Unity

Kant’s decision to compare his *Critique of Pure Reason* to Copernicus’ astronomical project is much referenced, but widely misunderstood. By taking a close look at what Copernicus’s famous revolution in astronomy actually involved, in this section I will argue that Kant uses the analogy to highlight his own use of the transcendental presupposition of the unity of reason. As we will see, both Kant and Copernicus grounded their respective disciplines by systematising the object of study for that discipline, and they both did this by showing how certain presuppositions were necessary for the possibility of the form of science that they wished to pursue. Through this analysis I also hope to defend the strong interpretation of Kant’s transcendental deduction of the unity of reason, which I sketched out above.

Kant’s reference to Copernicus comes in the B-Preface, where he draws an analogy between the project of the first Critique, and the so called “revolution” in the sciences brought about by Copernicus [Bxvi]:

---

11 This is also a retort to those who have claimed that Kant’s use of the Copernican analogy is meant to signal nothing more than the fact that both Kant and Copernicus relied on the use of presuppositions to advance their chosen field. See, for example, Hanson (1959, 278).
“We should then be proceeding precisely on the lines of Copernicus' primary hypothesis. Failing of satisfactory progress in explaining the movements of the heavenly bodies on the supposition that they all revolved round the spectator, he tried whether he might not have better success if he made the spectator to revolve and the stars to remain at rest. A similar experiment can be tried in metaphysics, as regards the intuition of objects. If intuition must conform to the constitution of the objects, I do not see how we could know anything of the latter a priori but if the object (as object of the senses) must conform to the constitution of our faculty of intuition, I have no difficulty in conceiving such a possibility” [Bxvi-xvii].

Our first response here should be to recognise that, when Kant draws this analogy, he is careful to begin by discussing our intuition of objects, and not our application of reason to the understanding. How then, would this claim have any connection to Kant’s introduction of the transcendental idea of the unity of reason, which explicitly deals not with intuition but with the unity of the concepts of the understanding? What I intend to show here is that the activity of intuition that Kant speaks of when drawing his analogy with Copernicus is made possible by a prior presupposition concerning the unity of reason, and that this too is analogous with the presuppositions that Copernicus made in his attempts to unify the science of astronomy. This reading is supported by the comments Kant makes immediately after the quotation given above, where he continues to show how his analysis of intuition cannot be complete without an accompanying analysis of both the understanding and reason, which he says “will furnish an excellent test of what we are adopting as our new method of thought” [Bxviii].

To understand what Kant means when he calls for a “similar experiment” to that of Copernicus, we will need to look at exactly what Copernicus did for the science of astronomy. It is
commonly understood that Copernicus’ brilliance consisted in the fact that he suggested a change of perspective for astronomers and that by looking at things from the perspective of a static sun, rather than from the perspective of a static earth, he removed the need for the mathematical complexity of the Ptolemaic methods and that he greatly improved the accuracy with which the location of the planets could be deduced. In fact, none of this is strictly true. The Copernican model did not improve the accuracy of astronomical predictions. In fact, the computations involved in predicting the locations of planets under the Copernican model yielded identical results to those produced by the standard Ptolemaic model. This is largely due to the fact that, while the cosmological picture of the solar system that resulted from Copernicus’ innovation marked a break from the Ptolemaic representation of the heavens, the mathematical methods used to make astronomical predictions were effectively identical. Consequently, as Thomas Kuhn points out, the methods of Copernicus and Ptolemy “could not be differentiated on consistency, nor by simplicity of computation” (Kuhn, 1977, 334). The technical situation is summarised by Matjaž Vesel when he writes that Copernican astronomy “relied on the traditional technical apparatus used by Ptolemaic astronomy, including eccentrics, deferents, epicycles, and epicycles upon epicycles” and that this theory “did not predict the positions of celestial bodies any better than that of Ptolemy” (Vesel, 2004, 167-168). This fact is worth keeping in mind when we try to understand why Kant drew an analogy between his own system and the Copernican model.

---

12 This fact is insisted upon by a number of scholars, yet widely disregarded in the popular account of the history of science: “Modern historians, making ample use of the advantage of hindsight, stress the revolutionary significance of the heliocentric system and the simplifications it had introduced. In fact, the actual computation of planetary positions follows exactly the ancient pattern and the results are the same” (Neugebauer, 1968, 103). Derek Price adds that the “angles, distances, and indeed all mathematical techniques are quite unaffected by any change from geostatic to heliostatic systems... Provided that suitable lines and planes of reference are maintained there is not a single mathematical technique or calculation which is peculiar to the geostatic or to the heliostatic system alone” (Price, 1969, 203).

13 Whether or not Copernicus was aware of the fact that his model did not improve the accuracy of astronomical predictions is unclear (Price, 1969, 198-199).
The question of whether the paradigm shift in astronomy instigated by Copernicus was the result of a simple change of perspective is more complicated, and by looking at it in detail we will begin to see what Kant and Copernicus share. It is worth pausing on this question of perspectival change because it is one which causes much confusion in the literature. It is often remarked, for example, that Kant’s supposed change of perspective moves in the opposite direction to that of Copernicus. While Copernicus showed that we should stop assuming that the subject is in the centre of the universe – so the story goes – Kant showed us that we need to put the subject back into the centre. This is a point made by Bertrand Russell: “Kant spoke of himself as having effected a ‘Copernican revolution’, but he would have been more accurate if he had spoken of a ‘Ptolemaic counter-revolution’ since he put Man back at the centre from which Copernicus had dethroned him” (Russell, 1948, 9). To see why Russell is mistaken, and why this visual metaphor of perspective is misleading, we must remember that Copernicus was concerned first with what we would call mathematical astronomy, and not with cosmology.

If Copernicus was a mathematician first, and if his astronomical model provided the same predictions as the mathematical model of Ptolemy, then why is he of any interest to the history of science, and what does Kant see in him? The simple answer is that, like Kant, Copernicus systematised a scientific discipline that was beginning to come apart at the seams, and he did so by effecting a change in the shared presuppositions of that discipline. Before Copernicus, the appearance of each of the planets was calculated by the use of a mathematical model that was specifically designed for that planet. In each case, a model could be drawn with the Earth at the centre and the planet in question moving along an orbit, while also deviating from this orbit according to its own epicycle and the relevant equant in each case (Price, 1969, 197-218).

14 This reflection is repeated so regularly it has become somewhat of a truism; Kemp-Smith (1999, 22), Guyer (2006, 50), and Schulting (2009, 43).
Each of these models was accurate enough, but they existed somewhat in isolation from one another. The distances used in these models to represent the relationships between the Earth and the planet in question were relative, so the different models could not easily be mapped onto one another. Furthermore, the discordancy of the different planetary models was exacerbated by the fact that the scale of the models for the inferior planets (Mercury and Venus) was difficult to match with the scale required for the superior ones (Mars, Jupiter, and Saturn). It is this proliferation of different planet-specific models, which led Kuhn to state that, under the Ptolemaic status quo, “[t]he astronomical tradition had become diffuse” and “no longer fully specified the techniques that an astronomer might employ in computing planetary position” (Kuhn, 1985, 139).

Copernicus’ innovation was to recognise that there was one shared value in the different planetary models. By using this value as a common point of reference, Copernicus was able to unify these various astronomical models into a system. As Derek Price writes: “It was already well known that, in addition to its own peculiar periodicity, each planetary theory contained an annual element; for the superior planets it was in the epicycle, for the inferior, in the deferent” (Price, 1969, 215). All that Copernicus had to do was to extract this common feature and hypothesise that it could correlate with the annual motion of the Earth and the Sun. The result was a complete systematic reorientation of astronomy, which nevertheless left all of the individual predictions completely intact: “In Ptolemaic theory all distances had been relative. In the new theory they were each related to the common element of the Sun-Earth system, and thereby related to each other” (Price, 1969, 216). Ultimately, the only difference between
Ptolemy and Copernicus was that the latter managed to overcome the multiplication of astronomical models and to combine these models into a systematic whole.\textsuperscript{15}

The fact that the predictions offered by the new system were identical to the predictions of the old system meant that no empirical evidence could be gathered to help astronomers decide between the old and the new ways. Given the fact that no such evidence could be weighed on either side, the decision came to nothing more than the question of which model should be presupposed. Ultimately, Copernicus argued, that our supposition should favour the systematic unity of the heliocentric model over the fragmented techniques required by the geocentric model, on the basis of the “coherence”, “harmony”, and “symmetry” of the former (Copernicus, 1992, 177).

With this recapitulation of the mathematical innovations of Copernicus in hand, we are now in a much stronger position to explain why Kant draws the analogy that he does.\textsuperscript{16} Much like Copernicus, Kant finds himself attempting to unify a discipline that has become fragmented. He begins the A-Preface lamenting the “intestine wars” that threaten to pull the science of reason apart and sets himself up as attempting to return order to the “complete anarchy” that has taken hold of the discipline [Aix]. Much like the methodology of Copernicus, Kant’s methodology involves finding something which is shared across the various models in need of

\textsuperscript{15} See Price: “[T]here never was such a thing, in this sense, as a Ptolemaic System. The complex mathematical theory is arranged so as to deal with each planet separately and individually. There is no single mathematical connection between these several models, only a general similarity in the methods used for each. Not before the work of Copernicus was any such mathematical link devised that welded the whole into a mathematical system rather than a cosmological one. This is, in fact the most important aspect of Copernican Theory—the invention of a mathematical planetary system, rather than the change from geocentric to heliocentric” (Price, 1969, 199-200).

\textsuperscript{16} It may be remarked here that the foregoing analysis of Copernicus draws on relatively recently scholarship, which would not have been available to Kant. On this account, we might expect Kant to make the same common errors that I have attributed to Russell, among others, above. While it is true that Kant cannot have been drawing on the same scholarship that I have referenced here, my claim is that by recognising what Kant and Copernicus actually share, we can offer a much more convincing analysis of the role that Copernicus plays in Kant’s self-reflective comments, and one which needn’t ascribe to Kant the great error of misnaming his theory Copernican rather than Ptolemaic.
unification and then using it to relativize everything around this common element. For Copernicus, this common element was the Sun-Earth relation, for Kant it is the subject-object orientation of all judgement. According to Kant’s narrative of the history of philosophy, the “dogmatists” and the “sceptics” both paid special attention to the relationship between the subject, and the associated powers of the subject’s mental faculties, and the nature of object, to which these faculties are applied [Aix]. By altering our philosophical presuppositions concerning the relationship between the subject’s powers of intuition and the object to which it applies, Kant sets out to unify the discipline of metaphysics around a shared systematic conception of the object of that science, ultimately incorporating aspects of both “dogmatism” and “scepticism” into this system. Speaking of the “experiment of pure reason” that is supposed to be analogous to Copernicus’ experiment with a new presupposition concerning the relation of the planets, Kant says that he is able to combine the two different conceptions of knowledge “namely, the knowledge of things as appearances, and the knowledge of things in themselves” into a complete “harmony” [Bxxi, footnote]. Thus, much like Copernicus’ revolution in astronomy, Kant’s reorientation in metaphysics involves a kind of methodological shift, which does not affect the results of individual enquiries. As we have seen, the perceived location of planets in the sky is not affected by Copernicus’ innovations: Whether the sun moves around the earth, or whether the earth spins on its axis while orbiting the sun, the observable phenomena are the same. Similarly, for Kant, whether the understanding must conform to objects, or whether objects of experience conform to our understanding, the phenomena will be the same. For example, whether reality is temporally ordered in itself, or whether intuition is necessarily temporal in its receptivity, either way we can know a priori that all objects of experience will be temporally extended. The difference in the two models is not in the phenomena themselves, but in the conceptual framework that explains the presentation of the
phenomena, and ultimately in our justification for empirical knowledge concerning the phenomena.

The relationship that this reading suggests between the faculties in Kant is as follows: while the intuition of the manifold by the faculty of sensibility can only become an object of knowledge because of the application of concepts carried out by the understanding, the systematic application of the understanding is only possible because reason systematises the understanding through the application of the transcendental idea of the unity of reason. The very shift in perspective which is supposed to be characteristic of both Copernicus and Kant is thus based on their shared attempts to unify a science through a prior presupposition of the unity of the concepts applied in that science.

The final point of analogy between Copernican astronomy and Kantian metaphysics thus concerns the way in which they both rely on a certain kind of presupposition. In the final section of this paper, I will sketch both the similarities and dissimilarities between Copernicus’ and Kant’s respective use of presuppositions and will demonstrate that Kant’s transcendental deduction of the unity of reason cannot coherently be understood as a heuristic device but must be understood as the statement of a constitutive law, by which we are compelled to presuppose not only the unity of reason itself, but also the real unity of nature.

4. Unity, Transcendental Presuppositions, and the Science of Metaphysics

If, as I have claimed, Copernicus’ proclamation of a heliocentric astronomical system left the predictions for planetary appearance completely unchanged, then it seems that Copernicus would not be able to supply any convincing arguments for why we should presuppose a
heliocentric model over a geocentric one. Put otherwise, it might seem as if Copernicus is relying on prefatory-presuppositions rather than suppling an argument in favour of a conclusory one. If it is also the case that Kant’s revolution in the field of metaphysics is analogous to the changes brought about in astronomy by Copernicus, then this would mean that Kant’s critical project is left in jeopardy of relying on prefatory-presuppositions too. To see why this is not the case, we must pay attention to the kind of arguments that support both Copernicus’ and Kant’s presuppositions. In both cases, we find that, on the condition that we desire to avoid self-contradiction, we do not have two possible presuppositions open to us after all.

The issue here concerns the necessary systematicity of a science. Kant writes that “systematic unity is what first raises ordinary knowledge to the rank of science” and that systematic unity is what “makes a system out of a mere aggregate of knowledge” [A832/B860]. Thus, the reasons that compel the astronomer to accept heliocentricity as a conclusory-presupposition concern the fact that without such a presupposition there could be no science of astronomy, because there would be no unified system of planetary motion, and thus no unified object for it to study. Kant makes a similar claim concerning his own Copernican revolution. He attempts to show that, if we presuppose an epistemological model in which our knowledge must conform to objects as they are in themselves, then it will be impossible for us to overcome the internal contradictions that pull apart the science of metaphysics. If, on the other hand, we presuppose that objects of knowledge must conform to our understanding, and that the concepts of our understanding form a systematic unity, then such contradictions vanish. As metaphysicians then, we are compelled to take the second option. Comparing his project to that of Copernicus, Kant writes the following:
“If, then, on the supposition that our empirical knowledge conforms to objects as things in themselves, we find that the unconditioned cannot be thought without contradiction, and that when, on the other hand, we suppose that our representation of things, as they are given to us, does not conform to these things as they are in themselves, but that these objects, as appearances, conform to our mode of representation, the contradiction vanishes… we are justified in concluding that what we at first assumed for the purposes of experiment is now definitely confirmed” [Bxx-xxi].

The point I would like to emphasise here, and one which will be articulated in more detail below, is that the latter presupposition is simultaneously a presupposition of the unity of the concepts of the understanding. To presuppose the former is to accept that our concepts do not form a systematic whole and thus to accept – willingly or not – that our application of these concepts will lead to contradictions. To orient our decision, and to ground metaphysics as a science, we must assume that our concepts are unified, and this is done via the transcendental presupposition of the unity of reason. In one and the same step, then, we must recognise the need to presuppose the unity of knowledge and the unity of the nature to which this knowledge corresponds. What I have been calling the transcendental presupposition of the unity of reason includes both of these claims. To refuse to make this presupposition would be to continue in a purely ad hoc fashion, and to give up on the possibility of a true science of reason.

The picture I have painted so far of the analogy between Copernicus and Kant, and the resulting defence of the transcendental presupposition of the unity of reason, requires clarification in at least two respects: First, one might ask, isn’t it the case that while Copernican mathematics said nothing of the actual location of the planets, this mathematical model did also result in an
alteration in the cosmological picture that Copernicus promoted? That is to say, didn’t Copernicus claim not only that we should assume the sun is in the centre of the solar system for the purposes of mathematical systems, but that the sun is actually in the centre of the solar system? If this is the case, then Copernican astronomy will be open to verification or falsification by empirical methods after all, and the search for the unity of a science of astronomy may need to give way to the material facts of the matter. I say this here because, if Kant’s move is supposed to be analogous, then perhaps it could also be the case that Kant’s metaphysical presuppositions could turn out to be in conflict with reality. The second issue, which is connected to the first, is that, regardless of empirical claims, Kant’s argument rests on the idea that we are compelled to create a unified science of reason. Even if Kant had shown that the only way to create a coherent metaphysics was to presuppose the unity of reason, he has not yet given us any reasons why we must take this as our aim. This latter issue is directly linked to the question discussed above of how to interpret Kant’s transcendental presupposition of the unity of reason. The “weak” interpretation, which took Kant’s presupposition to be a mere heuristic suggestion correlates with the possibility that a unified science of reason is not a necessity, but that we should choose to assume such a unification for pragmatic reasons. In order to defend the claim that Kant’s use of the Copernican analogy is intended to prepare the reader for a stronger interpretation of the transcendental deduction of the unity of reason, as it is presented later in the Appendix, I will now respond to each of these two issues in turn.

In answer to the first issue, concerning the empirically verifiable aspects of heliocentric cosmology, it is worth recognising that, while Copernicus sees a connection between mathematical astronomy and cosmology, it is still the former that conditions the latter. For Copernicus, “it is the mathematical astronomy which decides whether Earth moves (or not), not cosmology or physics” (Vesel, 2004, 186). In effect, Copernicus showed mathematically
that, if we want to think systematically about the movements of all of the planets, then we have no choice but to think of the sun in the centre. Then, having built a mathematical model for this system, it becomes possible to use this model as a representation of the actual location of the planets. As Price explains: “the Copernican invention of a planetary system was the first step in enabling the technical astronomer to conceive of the mathematical device of circles as representing the orbit of a physical planet in space” (Price, 1969, 201, my italics). On the one hand then, Copernicus’ heliocentric model was nothing but a hypothesis that would later be confirmed once the physical laws discovered by Newton and others could be applied to the empirical data, while on the other it was the ground for the possibility of this empirical study. This is something that Kant recognises when he makes his Copernican analogy. Acknowledging the work of figures such as Newton, Kant explains that the discovery of “the fundamental laws of the motions of the heavenly bodies gave established certainty to what Copernicus had at first assumed only as an hypothesis” [Bxxiii, footnote]. What is interesting here, is how Kant maps his own project onto this story. According to Kant’s analogy, the first Critique plays both of these roles at once. Kant sees himself as making the theoretical presupposition that would allow for the science of metaphysics and then also making the arguments which supply the necessary support for this hypothesis:

“The change in point of view, analogous to this hypothesis, which is expounded in the Critique, I put forward in this preface as an hypothesis only, in order to draw attention to the character of these first attempts at such a change, which are always hypothetical. But in the Critique itself it will be proved, apodeictically not hypothetically, from the nature of our representations of space and time and from the elementary concepts of the understanding” [Bxxiii, footnote].
What is striking about this comment is the fact that Kant specifically claims that it is the transcendental deductions of the forms of intuition and the categories of the understanding that will prove apodictically the truth of Kant’s presupposition. The transcendental deduction of the ideas of pure reason, on the other hand, is not seen as part of that proof. The reason for this, I would like to argue, is that this is the section of Kant’s Critique which is directly analogous to the work of Copernicus. Just as Copernicus first made a presupposition of the systematic unity of mathematical astronomy, which would later be proved correct by the application of the laws of motion to the empirical observation of the planets, Kant’s transcendental deduction of the ideas of pure reason leads him to presuppose the unity of reason, which will later be proved correct by the transcendental deductions of space and time and of the categories of the understanding. The order in which these arguments are presented in the first Critique should not mislead us here. While the transcendental deductions of the forms of intuition and the concepts of the understanding are presented first, they actually confirm, rather than lead up to, the transcendental deduction of the ideas of pure reason. Like Copernicus, Kant requires two stages to his proof, the first is a hypothetical presupposition, given in the preface and defended in the appendix, where it is considered from the point of view of the necessity of the systematic unity for any science, and the second is a proof which confirms the results of such a presupposition, given by the transcendental deductions of forms of intuition and the concepts of the understanding.

There remains here, of course, a “clear disanalogy” between Kant and Copernicus in one sense (Schulting, 2009, 59): while for Copernicus the second stage of the proof involves empirical data, for Kant, the second stage is provided by an apodictic proof in the form of a transcendental deduction. This should not be surprising to us, given the fact that “the typical scientist’s practice consists in positing falsifiable theses, whereas the philosopher does not
posit falsifiable theses but puts forward hypotheses with a view to apodictically proving principles that hold a priori and necessarily” (Schulting, 2009, 59). By now turning our attention to this dissimilarity between metaphysics and the other sciences we will be in a position to confront the second issue that I laid out above, namely to find the grounds for Kant’s assumption that we must attempt to found a unified science of reason, rather than simply admit the impossibility of metaphysics.

In the B-Preface, and by use of the Copernican analogy, Kant tells us that we must make a presupposition to allow for the possibility of a systematic metaphysics. What is more, he tells us that what we must presuppose is the unity of reason. This is precisely the claim that Kant repeats in the Appendix to the *Transcendental Dialectic*. As it is laid out here, it is the transcendental idea of the unity of reason that we must presuppose, so that reason is able to unify the concepts of the understanding, without which a systematic application of these concepts would be impossible in the acquisition of knowledge. Earlier on, I argued that Kant’s claim that we must presuppose the unity of reason should be interpreted in a strong sense and that the law which states this is in fact constitutive of the science of metaphysics. Now, with the similarities and the differences between the science of metaphysics and the science of astronomy in hand, it is possible for me to add more weight to my defence of this strong interpretation.

When presented with the choice of either presupposing the unity of reason in order to allow for the possibility of the science of metaphysics, or presupposing a fragmented reason and thus making metaphysics impossible, what is it that compels us to take the former route? Is it simply

---

17 For a defence of Kant’s non-empirical experimental method, which rebuts Popper’s criticism of Kant on the grounds that his philosophy is not falsifiable see Schulting (2009, 48-49 footnote).
the fact that, for pragmatic reasons, it is more useful to imagine the unity of reason? I think not. The explanation for this fact is given when we recognise that, for Kant, it is not possible for us to avoid metaphysics. Reason has a natural tendency to ask questions which cannot be answered empirically. For this reason, it “is burdened by questions which, as prescribed by the very nature of reason itself, it is not able to ignore” [Avii]. Consequently, Kant will claim that “in a certain sense” metaphysics “is to be looked upon as given; that is to say, metaphysics actually exists, if not as a science, yet still as natural disposition (metaphysica naturalis)” [B21]. This distinguishes the science of metaphysics from the other sciences, such as astronomy. While it could turn out to be the case that there is no systematic object of study for astronomy, from which it would follow that there simply is no science of astronomy, this cannot be an option for metaphysics. The discipline of metaphysics is “a completely isolated speculative science of reason” which we must grapple with, whether we want to or not [Bxiv]. Effectively, Kant claims that due to the natural tendencies of reason, we will have a metaphysics of some sort. The only question that remains is whether we make a science out of that metaphysics by applying reason’s own rules to itself, and thus presupposing the unity of reason, or whether we allow reason to fall into a state of contradiction in which dogmatism and scepticism remain eternally irreconcilable. For Kant, “the critique of reason, in the end, necessarily leads to scientific knowledge; while its dogmatic employment, on the other hand, lands us in dogmatic assertions to which other assertions, equally specious, can always be opposed that is, in scepticism” [B23].

When Kant puts forward the transcendental presupposition of the unity of reason, by which “we have no option save to presuppose the systematic unity of nature” [A651/B679], he is stating a regulative law which is constitutive of the science of reason known as metaphysics. It is not merely a heuristic that we should utilise without taking what follows from it seriously.
Unless “a systematic unity is a priori assumed to be necessarily inherent in the objects”, the natural tendencies of reason will lead us to a self-contradictory and unscientific metaphysics [A651/B679]. This is what the Copernican analogy in the B-Preface is designed to prepare us for. Thus, the transcendental idea of the unity of reason is a conclusory-presupposition, first hypothesised in the B-Preface, defended by the transcendental deduction of the ideas of pure reason offered in the Appendix, and then proved by showing that the transcendental deductions of space, time, and the concepts of the understanding would not be possible without it. The transcendental deduction of the presupposition of the unity of reason “will complete the critical work of pure reason” because it is constitutive of the science of metaphysics itself [A670/B698].

**Bibliography:**


