1. Introduction

Writing on “Kant on method” is an intimidating task. After all, as Kant suggests in the B Preface to the first Critique, the critical philosophy is in some sense “a treatise on the method” of metaphysics or philosophy (Bxxii). Thus, it is appropriate that each of the Critiques concludes by explicitly turning to methodological questions. Given this, a complete presentation of Kant’s views about philosophical methodology would require at least a treatment of all three Critiques. But the topic of “Kant on method” is also daunting for other reasons. For while Kant certainly does have certain general methodological views, his conception of the proper methodology for different areas of inquiry varies considerably with the area of inquiry in question – so that different methods are appropriate to (say) mathematics or physical science or anthropology or foundational moral philosophy.

This methodological variation can be traced, in considerable measure, to two other axes of variation between these areas. First, there is variation in the resources that are available to us as human beings to answer questions in this or that domain. For example, as mathematicians we have available to us resources – like construction in a priori intuition – that are not available to us, at least in the same way, as philosophers. Similarly, as practical or moral philosophers, we have resources – like those provided by our consciousness of the moral law as the principle of pure practical reason – that are not available to us in purely theoretical or speculative forms of inquiry. Moreover, and this is the second axis of variation, Kant also believes that there are important differences between the aims or goals of different forms of inquiry. And these differences matter for proper methodology in those areas. So, for example, Kant’s conceives of philosophical anthropology as having “pragmatic” aims that are much less cognitively demanding than the aims that are associated with foundational theoretical or practical philosophy. Ultimately, then, Kant presents us with a quite varied picture of proper methodology for different areas of inquiry. But, nonetheless, for Kant, there are certain general observations we can make about both (i) the aims of human inquiry and (ii) the resources

---

2 For Kant on the relationship between the methodology of mathematics and philosophy, see: (Dunlop 2012; Heis 2014; Sutherland 2010; Nunez 2014).
3 For the methodological significance of the Fact of Reason, see (Schafer 2019a; forthcoming). For further discussion, see (Kain 2010; Kleingeld 2010; Schönecker 2013; Ware 2014; Tenenbaum 2012). Other relevant factors here concern the proper role of analogy and aesthetic judgment in different areas of inquiry. See (Breitenbach 2014; Kohl 2015).
4 For more on Kant’s anthropological works, see (Cohen 2009).
available to us in our pursuit of those aims. And understanding these general features of human inquiry is crucial for understanding Kant’s conception of the proper methodology for philosophy – which, for Kant, provides the foundation for all inquiry, at least insofar as it can be given a rationally satisfying form.

In section 2, I offer a brief characterization of how Kant conceives of the aims of human inquiry – focusing on the idea that inquiry ideally aims at, not just cognition (Erkenntnis), but rather the more demanding cognitive achievements that Kant labels insight (Einssehen) and comprehension (Begriffen). Indeed, as we will see, talk of “method” in a strict sense is appropriate for Kant only with respect to forms of inquiry that aim at the sort of systematic unity which makes “insight” and “comprehension” possible. In section 3, I will consider the implications of this picture of inquiry for the proper practice of philosophy. Here I will emphasize Kant’s distinction between critical and doctrinal (or metaphysical) phases of philosophical inquiry – focusing on how this first, critical phase plays both a negative and a positive role with respect to the second, doctrinal phase – both “disciplining” it by setting its proper boundaries and establishing the foundations it requires. In particular, I will explain how the critical phase of philosophy is meant, for Kant, to uncover and clarify the organizing principles or ideas that are required if a genuinely systematic or “scientific” form of doctrinal philosophy or metaphysics is to be possible. In section 4, I will argue that this positive role for critique is possible, according to Kant, only insofar as philosophy follows what I call a “capacities-first” methodology – that is, a methodology that treats certain cognitive capacities (such as reason) and their self-conscious activities as fundamental (in both a cognitive or epistemic sense and in an explanatory or metaphysical sense) for the sort of philosophy human beings are capable of. It is this methodology, I will argue, that allows Kant introduce the first principles that philosophy in its doctrinal phase requires, and to do so in a manner that is neither arbitrary nor (at least obviously) incompatible with the application of Kant’s critical restrictions to his own philosophy. Finally, in the remainder of section 4 and in section 5, I will briefly discuss some of the implications of this methodological picture – including its implications for the methodological significance of self-consciousness and regressive or “transcendental” arguments, the relationship between it and Kant’s distinction between analytic and synthetic methods in philosophy, the role of “transcendental deductions” on this account, and its connections with Kant’s famous claim that the philosopher – unlike the mathematician – must be a legislator, and not an artist, of reason.

2. The Aims of Rational Inquiry: Cognition, Knowledge, and Comprehension

5 For more on this, see (Schafer forthcoming; forthcomingd). Compare (Tolley 2017; Breitenbach 2018; Zinkin 2017; Proops 2021). For systematicity in post-Kantian philosophy, see (Franks 2005; Förster 2011).
6 See (Schafer 2019a; 2019b; 2018). For some broadly similar ideas, see, for example, (Kern 2006; 2017; Land 2018; Merritt 2018; Engstrom 2016).
7 The central role of capacities in Kant’s methodology can be seen in his definition of a “canon” as “the sum total of the a priori principles of the correct use of certain cognitive faculties in general.” (A796/B824)
Let's begin with how Kant conceives of the goals of human inquiry. Here three cognitive achievements are particularly important – namely, what Kant calls cognition (Erkenntnis), knowledge (Wissen), and comprehension (Begreifen). Although there is no space to discuss the relationship between these in detail, in order to understand Kant’s philosophical method, it is crucial to begin by saying a bit about them. For in order to understand Kant’s method, we need to understand what he hopes to achieve through it. And to understand this, we need to understand how Kant thinks about the cognitive achievements at which rational inquiry in general aims.

In the first instance, Kant conceives of inquiry as aiming at what he calls “cognition”. To cognize something, in the sense relevant here, is to form an “objective representation” of it “with consciousness” (B376–7; cf. JL 9:65, BL 24:133-4). That is, cognition is the cognitive achievement through which we become conscious of objective features of reality – as opposed to merely being conscious of our own subjective states. In this way, to have cognition in this sense is to be conscious of something that imposes a real or “material” standard of correctness on our representations of it – that is, a standard of correctness that goes beyond the merely “formal” or logical standards of correctness that are internal to any (conceptual) representation as such.

Crucially, such cognition may be either theoretical – that is, cognition of what is – or practical – that is, cognition of what ought to be. As Kant understands this distinction, theoretical cognition aims to accurately represent an independently existing object or state of affairs – while practical cognition aims to make it the case that the objects it represents exist. Thus, theoretical cognition is fundamentally receptive for Kant, while practical cognition is fundamentally productive or active. But, in both cases, cognition is only possible for Kant insofar as there is a non-accidental relationship between our conscious representation of some object and that object – and this relationship must itself be available to consciousness in some sense. For it is only insofar as a cognition makes its relationship with its object available to consciousness that it allows us to become conscious of this object as something objective in the sense laid out above. Thus, we can achieve cognition in this sense only insofar we have conscious access to (i) an object (either as what is or as what ought to be), (ii) a representation of that object, and (iii) a non-accidental relationship between that object and that

---

8 Beginning with Smit’s very important work (Smit 1999; 2000; 2009; 2019), there has been an explosion of work on Kant’s conception of cognition. For my views, see (Schafer forthcoming; forthcomingd). But also see (Watkins and Willaschek 2017; Tolley 2017; Chignell 2014) for important discussion.

9 As Matt Boyle has noted to me, when we are dealing with our own empirical self-consciousness (in inner sense), it becomes less clear how exactly to distinguish objective from merely subjective consciousness from one another. For more on this, see Boyle’s contribution to this volume.

10 See (Schafer forthcoming; forthcomingd).

11 Here I draw on the important work of (Engstrom 2002; 2006; 2009). See also (Merritt 2018). For my views on this issue, see [REDACTED]. For other views of these issues, see (Neiman 1994; O’Neill 1989; Cohen 2014; Mudd 2017; Guyer 2006; 2019)
representation, which will be receptive in the case of theoretical cognition and productive or active in the case of practical cognition.\footnote{Compare (Tolley 2017). Lest this seem too demanding, it is important to stress that having conscious access to these three aspects of cognition does not require us to possess the concepts required to form explicit discursive thoughts about them – although it does put in me in a position to form such concepts through reflection upon my cognitions. For more on the sort of self-consciousness here, see Boyle’s contribution to this volume.}

In this way, for Kant, genuine inquiry is distinguished from the mere subjective play of our faculties precisely insofar as inquiry aims to get at \textit{something objective} over and above the subject’s fleeting subjective states. More precisely, in aiming at cognition, inquiry aims to bring this “objective something” (be it theoretical or practical) to consciousness. As I’ve argued elsewhere, this in turn imposes two further, important constraints on cognition.\footnote{See (Schafer forthcoming; forthcomingd).} First, in order to cognize something in this sense, we must be conscious of that object of cognition as something that is really (as opposed to merely logically) possible.\footnote{Compare (Chignell 2014).} And second, we can only be said to have cognition of an object insofar as our representation of it is sufficiently determinate to provide us with a meaningful standard of material correctness (be this standard theoretical or practical).\footnote{Compare (Ameriks 2003). For example, it is the “indeterminacy” of the ideas of pure speculative reason that primarily explains their inability to provide us with cognition of their objects. Compare (Walden 2019).}

For Kant, the characteristic aim of the understanding, as the faculty for discursive \textit{cognition}, is to determine the objects of our representations so as to meet these two constraints. But, for Kant, inquiry does not merely aim at \textit{bare} cognition. For the ultimate aim of human inquiry is not mere cognition (\textit{Erkenntnis}) as such, but rather the further perfections of cognition that Kant associates with what he calls “insight” (\textit{Einsehen}) and “comprehension” (\textit{Begreifen}).\footnote{\[REDACTED\].} In essence, these “higher” forms of cognition go beyond \textit{mere} cognition insofar as they satisfy – not just the understanding’s need for determinate cognition of objects – but also reason’s need for a systematic grasp of \textit{how} and \textit{why} things are (or ought to be) as one cognizes.\footnote{In claiming that these perfections of cognition go beyond “mere cognition” I don’t mean to deny that there is a sense in which all cognition, for Kant, aims at this sort of systematic cognition.} Thus, insight and comprehension go beyond mere cognition insofar as they require not merely the “coordination” of marks that is required to distinguish one thing from another, but also the “subordination” of marks that is required in order to grasp explanatory relationships. (JL 9:134) In this way, insight and comprehension in Kant’s sense go beyond piecemeal cognition of particular objects (be it theoretical or practical) because they place this cognition into a more general \textit{system} – a system that allows us to explain and justify what we cognize (either theoretically or practically):
If we survey the cognitions of our understanding in their entire range, then we find that what reason quite uniquely prescribes and seeks to bring about concerning it is the **systematic** in cognition, i.e., its interconnection based on one principle. This unity of reason always presupposes an idea, namely that of the form of a whole of cognition, which precedes the determinate cognition of the parts and contains the conditions for determining *a priori* the place of each part and its relation to the others. Accordingly, this idea postulates complete unity of the understanding's cognition, through which this cognition comes to be not merely a contingent aggregate but a system interconnected in accordance with necessary laws. (A645/B673; cf. A680/B708, A738-9/B765-6)

Once again, it is important to stress that “comprehension” in this broad sense comes in both a theoretical and a practical form – although Kant often prefers to use “wisdom” (*Weisheit*) to refer to its practical variant. As I've argued elsewhere, this is crucial for understanding the **unity of reason** in its theoretical and practical uses.\(^{18}\)

For given this, we can characterize the aims of reason in general in terms of the drive for insight and comprehension – and then distinguish theoretical from practical reason in terms of whether the comprehension in question is theoretical or practical. But what matters here is that Kant, drawing on the traditional notion of *scientia*, conceives of this sort of **systematic understanding** as what distinguishes genuinely “scientific” or “rational” cognition from ordinary “popular” cognition: “systematic unity is that which first makes ordinary cognition into a science, i.e., makes a system out of a mere aggregate of it” (A832/B860). Thus, given that reason aims at systematic comprehension, a body of cognition will only be *fully* rationally satisfactory for Kant insofar as it possesses the systematic unity that makes comprehension possible. In this sense, reason will be satisfied by a body of cognitions only insofar as it takes the form of a “science” (*Wissenschaft*).\(^{19}\)

Moreover, as the term “*Wissenschaft*” suggests, the sort of systematic unity that is characteristic of comprehension for Kant is also closely associated with what Kant calls “knowledge” (*Wissen*) – that is, with taking something to be true on grounds that are both “subjectively and objectively sufficient”.\(^{20}\) In order to have such grounds, at least two things are required for Kant. First, we must connect our representations together in a systematic manner so that one representation is available as a potential ground for others. And, second, at least some of these representations must provide us with the consciousness of objects that is characteristic of cognition, a consciousness that allows them to serve as grounds for assent that are sufficient in both a subjective and an objective sense. Thus, although, strictly speaking, comprehension and knowledge

---

\(^{18}\) [REDACTED]. For related ideas, see (Merritt 2018; Guyer 2019). Contrast this approach to the “unity of reason” with the more “pragmatic” line taken by (O’Neill 1989; Neiman 1994; Cohen 2014; Mudd 2017).

\(^{19}\) As Kant stresses, this sets a very high standard, which is beyond the cognitive capacities of human beings in many domains of inquiry.

\(^{20}\) For more on “taking to be true”, see (Chignell 2007a; 2007b; 2014). For a contemporary variant on this account, see (Schroeder 2015b; 2015a).
involve perfections of cognition along two distinct axes of perfection, they remain closely connected for Kant. For Kant, knowledge is the highest perfection of the modality of cognition – roughly, the manner in which it involves a “holding to be true” (Fürwahrhalten) – while comprehension is the highest perfection of cognition in relation to its “objective content”. But, while these perfections are logically distinct, to seek one of them is also to seek the other. And so, in searching for systematic comprehension, reason is also searching for a system of grounds of the sort that knowledge (Wissen) requires.

Now, as Kant stresses, the “scientific” or systematic form of cognition that is at issue here is only possible insofar as cognition proceeds in accordance with a “method”:

Cognition, as a science, must be arranged in accordance with a method. For science is a whole of cognition as a system, and not merely as an aggregate. It therefore requires systematic cognition, hence one composed in accordance with rules on which one has reflected. (JL 9:139)

Indeed, it is only here, with a systematic form of cognition, that talk of a “method” becomes strictly speaking appropriate for Kant: “for popular cognition needs a manner but science a method, i.e. a procedure in accordance with principles of reason by which alone the manifold of a cognition can become a system” (CPrR 5:151; cf. DWL 24:779). That is, talk of “method” is only (strictly speaking) appropriate for Kant insofar as the method in question proceeds under principles that can give systematic unity to a body of cognitions: “If something is to be called a method, it must be a procedure in accordance with principles” (A855/B883). In this sense, any genuine method, for Kant, is associated with a form of systematic unity. It is in this spirit that Kant writes that, “the universal doctrine of method ... [deals] with the form of a science in general, or with the ways of acting so as to connect the manifold of cognition in a science” (JL 9:139). And similarly: “By the transcendental doctrine of method, therefore, I understand the determination of the formal conditions of a complete system of pure reason” (A708/B736).

In this way, for Kant, “The doctrine of method is supposed to expound the way for us to attain the perfection of cognition” (JL 9:140). Once again, the highest perfection of cognition is, for Kant, to be found in the sort of systematic understanding that is characteristic of insight, comprehension, and scientific knowledge. So, while there will of course be “methods” in some sense that aim at less demanding cognitive achievements (at mere cognition or mere belief), it is distinctive of “methods” in a strict sense that they characterize the forms of thought necessary to achieve systematic unity and, with it, insight and comprehension.
 Crucially, this means that such “methods” must provide us with both (i) the consciousness of determinate objects that is characteristic of cognition and (ii) the consciousness of systematic unity that insight and comprehension require. As a result, any genuine method – of either a theoretical or practical sort – can be thought of as aiming to simultaneously meet these two demands. (Compare JL 9:140) This, it is important to stress, is by no means an easy task for Kant. For, at least according to Kant, these two basic constraints on systematic comprehension of things often push finite creatures like us in opposed directions. For example, as Kant discusses in the Transcendental Dialectic, reason’s drive for systematicity pushes us to postulate an unconditioned object that can ground a complete explanatory system of theoretical cognitions. But, at least in a theoretical context, this act of postulation is unable to satisfy the constraints that must be met for us to have cognition of a determinate object. So, at least in a theoretical context, we cannot fully achieve the sort of systematic cognition of objects that reason craves. And while we can do better on this score in a practical context – insofar as practical reason provides with a practical consciousness of the moral law as the first principle of practical philosophy – there too practical reason (at least in human beings) faces the endless task of trying to make this abstract consciousness of the first principle of practical reason fully determinate in the manner that complete practical comprehension (or wisdom) and action requires. Thus, here too the task of trying to simultaneously meet these two constraints is a never-ending one for finite human beings like us.

The need for a method to satisfy both of these constraints is also closely related to one of Kant’s most famous methodological distinctions: the distinction between analytic (or regressive) and synthetic (or progressive) methods which he uses to distinguish his more popular philosophical works, like the Prolegomena, from the three Critiques. The first of these methods is characterized by a regress from what is conditioned in some domain to the more fundamental principles that condition and ground these conditioned objects and cognitions. Thus, it is characteristic of the first, “analytic” method that it begins with our cognition of determinate objects – be this theoretical, as in the case of empirical science, or practical, as in the case of ordinary moral thought – and then searches for the more universal principles that ground these particular cognitions. The second “synthetic” method, on the other hand, begins with the relevant principles and then progresses “downward” from them to give a full and systematic account of the relevant domain. Thus, while the analytic method can be thought of as beginning with the determinate objects of ordinary theoretical and practical cognition, and then trying to “solve for” the universal principles that give systematic unity to those objects – the synthetic method proceeds in the opposite fashion, beginning with the relevant principles and then attempting to “solve for” the determinate particulars that those principles explain.

23 (Gava 2015).
As this should make clear, the existence of these two methods follows from Kant’s general understanding of the proper aims of human inquiry. Thus, it is no surprise that both have a role to play within both theoretical and practical philosophy for Kant. But nonetheless, there is also a sense in which Kant associates the difference between these methods with the directions of thought that are characteristic of theoretical and practical reasoning respectively. For example, as fundamentally receptive in its relationship to objects, theoretical cognition (at least in human beings) must generally begin with the determinate objects that are given it by intuition. Only then can it attempt to discover the general principles that ground these objects and their properties. In this way, human theoretical cognition generally functions in a manner that is closer to the analytic or regressive method. On the other hand, practical reasoning (at least in its moral form) begins, for Kant, with a grasp of the ultimate practical principle – that is, the moral law. Thus, the characteristic task facing us as practical reasoners is less a matter of uncovering this principle as an abstract rule, and more a matter of determining which determinate actions it demands of us in our concrete circumstances. In this way, the characteristic direction of practical thought lines up more naturally with the synthetic or progressive method as opposed to the analytic or regressive one.

Nonetheless, both methods continue to have an important role to play in both theoretical and practical philosophy. After all, if our goal is to uncover the principles at work in any area of cognition or thought, starting with ordinary cognition of the relevant sort – it is the analytic method we should apply to that thought (even if that thought itself functions in a manner that is more akin to the synthetic method). On the other hand, in order to develop a systematic picture of the a priori principles at work in any area of cognition, theoretical or practical, we must show how those principles ground the relevant forms of a priori cognition in a systematic manner – something which the synthetic method does far better than the analytic. But nonetheless, there is a sense in which the analytic method (as regressive) is essentially connected, for Kant, with the idea of theoretical cognition as receptive – while the synthetic method (as progressing from principles) is essentially connected with the idea of practical or moral cognition. Thus, while it would be a mistake to view Kant’s synthetic account of the a priori elements of theoretical cognition in the first Critique as involving any sort of reduction of theoretical cognition to practical cognition – this account nonetheless rests on the recognition that there are fundamental respects in which a priori theoretical cognition is akin to practical or moral cognition.

3. Philosophy in its Critical and Doctrinal Phases

These points apply in some form to any area of human inquiry – although our ability to achieve insight and comprehension (or even cognition) will vary with the domain at issue. But what is most important for present purposes is how they apply to philosophy itself. Given that reason aims at systematic comprehension,
philosophy will only be satisfactory to reason insofar as it takes on this form. And for philosophy to do so, its cognitions must be unified by some idea or principle that characterizes their *unity as a system*.

Philosophy in *sensu scholastico* involves two things, (1.) A sufficient supply of cognitions of reason. (2.) A correct connection of these, or a system. For a system is the connection of many cognitions in accordance with an idea. (VL 24:799)

The two must be united; for without cognitions one will never become a philosopher, but cognitions alone will never constitute the philosopher either, unless there is in addition a purposive combination of all cognitions and skills in a unity, and an insight into their agreement with the highest ends of human reason. (JL 9:25)

At least according to Kant, the only alternative to such a “systematic philosophy” is what Kant calls the “naturalistic” method in philosophy. (A855/B883) Not surprisingly given his conception of reason as aiming at systematic comprehension, Kant takes such “philosophical naturalism” to be a disguised form of “misology” or disdain for reason and its aims. Thus, for Kant, for philosophy to take on a rationally satisfying form, both philosophical cognition *and* philosophical methodology must be given systematic unity by some overarching principle or idea.

This raises what is, in some ways, the fundamental question about the possibility of rational philosophy for Kant and the German idealists – namely, how we can locate a principle or idea that can play this foundational role – and, in particular, how we can do so in a non-arbitrary and rationally satisfying fashion.27 Putting aside those “philosophical naturalists” who do not concern themselves with the need for a genuine philosophical system, Kant divides previous attempts to respond to this challenge into two broad camps – “philosophical dogmatists” and “philosophical skeptics”. For Kant, “dogmatic” philosophy begins from the idea that “the method for obtaining certainty” in philosophy is fundamentally the same as it is in mathematics (A713/B741). In particular, the philosophical dogmatist assumes that philosophy has available to it *immediately certain and self-evident synthetic principles* (or “axioms”) which can play the role of philosophical first principles (A732/B760).28 But while such “axioms” are, for Kant, available to us in mathematics (which can rely on *a priori* intuition) – at least in theoretical philosophy, there is nothing that can play this role in a philosophical context.29 Thus, the model of mathematical cognition is a dangerously misleading one for the philosopher. At least in general, a

---

27 For an overview of this, see (Förster 2011; Horstmann 2012).
28 Wolff is quite explicit about this: “The rules of philosophical method are the same as the rules of mathematical method.” (Disc, 139) For the relationship between Wolff and Kant, see (Gava 2019; Dyck 2011; Chance 2018).
29 “Philosophical cognition is rational cognition from concepts, mathematical cognition that from the construction of concepts.” (A713/B741)
**priori** synthetic principles in philosophy cannot be justified by an appeal to pure intuition in the manner they are in mathematics. Rather, as Kant stresses, such principles require a “thorough deduction” which “justifies their authority” (A734/B762). It is the lack of such a deduction that, according to Kant, leads philosophers in the “skeptical” camp (including Hume) to call into question whether these synthetic *a priori* principles have any legitimacy at all.30

Against the background of this opposition between “dogmatism” and “skepticism”, Kant believes that only his “critical” method allows us to generate the *a priori* principles and ideas that philosophy requires without simply assuming them as “immediately certain” in the manner characteristic of philosophical dogmatism. As such, the proper methodology for philosophy involves two fundamental stages. First, a “critical” stage in which we arrive at these principles through a critique of our basic cognitive faculties. And second, a “doctrinal” or “metaphysical” stage in which we *use* these principles to construct a systematic account of the objects of theoretical and practical cognition – that is, of nature and morality:

The critique of the faculties of cognition with regard to what they can accomplish *a priori* has, strictly speaking, no domain with regard to objects, because it is not a doctrine, but only has to investigate whether and how a doctrine is possible through it given the way it is situated with respect to our faculties (CPJ 5:176; cf. 20:195).

Now the philosophy of pure reason is either propaedeutic (preparation), which investigates the faculty of reason in regard to all pure *a priori* cognition, and is called *critique*, or, second, the system of pure reason (science), the whole (true as well as apparent) philosophical cognition from pure reason in systematic interconnection, and is called *metaphysics* ... Metaphysics is divided into the metaphysics of the speculative and the practical use of pure reason, and is therefore either *metaphysics of nature* or *metaphysics of morals*. (A841/B869)

In thinking about these two phases of philosophy, it is important to stress that the role of the first is both negative and positive for Kant. To begin with the negative, critique in Kant’s sense provides our cognitive faculties with a source of “discipline” by determining their boundaries in a principled fashion. This, of course, is an aim that Kant shares with many other early modern philosophers. But Kant takes himself to go beyond these philosophers (e.g., Locke and Hume) insofar as his philosophy provides a genuine “critique of pure reason” as opposed to a mere “censorship” of it. That is, his philosophy, “subjects to evaluation not the *facta* of reason but reason itself, as concerns the entire capacity and its suitability for pure *a priori* cognitions.” As

---

30 For Kant’s view of Hume, see (Schafer, n.d.).
such, a genuine critique of pure reason establishes, “not merely limits but rather the determinate boundaries” for reason, boundaries that are “not merely suspected but are proved from principles” (A761/B789).

Given this, even in its negative role, a genuine critique of reason is possible only insofar as its claims about the boundaries of human cognition are grounded in a principled understanding of the nature of reason and our other rational capacities. In this sense, as we will see, genuine critique differs from mere censorship insofar as it is grounded in reason’s own self-understanding—a self-understanding which provides critique with the first principles it needs to have systematic unity. In this way, as Kant writes, the negative phase of critique operates by “revealing the deceptions of a reason that misjudges its own boundaries and of bringing the self-conceit of speculation back to modest but thorough self-cognition [Selbsterkenntnis] by means of sufficient illumination of our concepts”. (A735/B763)

Thus, a genuine critique of pure reason in Kant’s sense is only possible insofar as that critique provides itself with the a priori principles that a systematic account of the nature of our cognitive powers requires. So even the negative aspect of critique requires that it provide philosophy with philosophical principles in some sense. But the positive role of philosophy’s first, critical phase does not end there. For, as we will see, the essence of Kant’s “Copernican revolution” in philosophy lies in the idea that the very same principles, which explain the activity of our cognitive faculties, can also serve as the foundation for the second, “metaphysical” or “doctrinal” phase of philosophy as well. Thus, for Kant, we arrive at the foundational principles required by a “metaphysics of nature” and a “metaphysics of morals” precisely through a principled critique of our cognitive faculties. In this sense, the most important role of critique is not to give us a systematic understanding of the boundaries of our cognitive faculties, but rather to give us the principles that philosophy requires in order to construct a systematic picture of the natural and moral world(s). So, as Kant writes looking back on the first two Critiques from the third, just as the “critique of pure theoretical reason ... yielded the laws of nature”, the “critique of practical reason [yielded] the law of freedom,” thereby securing the foundational principles for these two areas of philosophy. (CPJ 20:202, compare A707/B735)

4. The Critical Philosophy as the Self-Consciousness of Reason

The question, then, is how the critical philosophy can provide us with such principles without sliding into a form of philosophical dogmatism. If what I have been saying so far is correct, Kant’s answer to this is that such principles can be provided only by the self-conscious rational faculties that are, for Kant, constitutive of human rationality. As I’ve argued elsewhere in more detail, such capacities are uniquely suited to play this role, for Kant, because they provide us with principles that are neither trivial nor rationally arbitrary. In this way, it is this system of self-conscious rational capacities (and their own internal principles) which, for Kant, grounds
the possibility of a systematic form of doctrinal philosophy or metaphysics – be this a metaphysics of nature in the theoretical domain, or a metaphysics of morals in the practical. In this sense, Kant’s critical philosophy may be thought of as a “capacities-first philosophy”.\(^{31}\)

To understand how the nature of these rational capacities makes philosophy possible for Kant, I want to briefly run through a few of their important features. First, in virtue of their self-conscious character, their activities must (at the very least) be accessible to conscious reflection. Just what this requires is, of course, a vexed question among Kant scholars – and it is one that I will not attempt to answer here.\(^{32}\) But it is plain that Kant believes that capacities like reason or the understanding can play the role they do in our cognitive lives only insofar they provide us with a consciousness of at least some of their acts. For example, as Kant says of the acts of the understanding in the A Deduction: “the mind could not possibly think of the identity of itself in the manifoldness of its representations, and indeed think this \textit{a priori}, if it did not have before its eyes the identity of its action, which subjects all synthesis of apprehension (which is empirical) to a transcendental unity, and first makes possible their connection in accordance with \textit{a priori} rules.” (A108) In this way, the possibility of a consciousness of the activities of these capacities is, for Kant, a necessary condition on their role in making human rationality possible.

Second, this consciousness is not merely a consciousness of the individual acts of these capacities. It is also, at least implicitly, a consciousness of the principles that govern these acts. For Kant, the idea that a self-conscious capacity must provide us with an (implicit) consciousness of its governing principles follows from the fact that it is these principles which give the acts of this capacity their characteristic form. Thus, to be conscious of these acts requires at least an implicit consciousness of the principles that govern them. We can make this general observation about the hylomorphic character of these acts more precise by remembering from above that, for Kant, genuine (theoretical or practical) cognition (\textit{Erkenntnis}) involves an implicit consciousness of the material standards of correctness that govern it. Thus, in order to function as capacities for cognition, our self-consciousness capacities must provide us, not merely with a consciousness of their acts, but also with a consciousness of the principles that govern the correctness of these acts.

As this indicates, Kant’s account of this sort of rational self-consciousness can be seen as anticipating

\(^{31}\) Here I take partial issue with (Lu-Adler 2018)’s claim that the foundations of Kant’s philosophy rest on “an eclectic procedure with distinctively Kantian characteristics.” (23) See (Schafer 2019a; n.d.). Compare again (Kern 2006; Land 2018).

\(^{32}\) For recent discussion see (Longuenesse, 2017; Boyle 2011; forthcoming; Land 2018; Smit 2019). Note that the role that this active self-consciousness plays in Kant is quite different from how philosophers like Wolff treated \textit{theoretical psychology} as foundational for logic.
many of elements of the contemporary discussion of whether rational activities like inference must be governed by some sort of “taking condition”.

Similarly, my focus on the foundational role of active, first-personal self-consciousness in Kant’s philosophy is in fundamental agreement with Schapiro’s recent claim that “the philosophical need to which critique is a response is a need for … a ‘guiding conception’ of the activities that we undertake.” But, as I read him, Kant believes that such a “guiding conception” is already implicit in the nature of the self-conscious activities we are engaged in and the capacities of which those activities are acts. Thus, to achieve a “guiding conception” of these activities, we need only to better understand what those activities and capacities are – from an active, first-personal, self-conscious perspective.

Thus, as Kant writes in the second Critique.

We can become aware of practical laws just as we are aware of pure theoretical principles, by attending to the necessity with which reason prescribes them to us and to the setting aside of all empirical conditions to which reason directs us. The concept of a pure will arises from the first, as a consciousness of a pure understanding arises from the latter. (CPrR 5:30, *my emphasis*)

As this indicates, to become conscious of the principles that are constitutive of practical reason and the understanding respectively, we need only to reflect on the activity of these faculties, abstracting away from the ways in which our empirical, sensible nature conditions it. It is this sort of reflection on the fundamental aims and principles of our faculties that, for Kant, provides philosophy with the basic principles it requires. The result, Kant believes, is “a system that takes no foundation as given except reason itself, and that therefore tries to develop cognition out of its original seeds without relying on any fact whatever” (P 4:274). Or, as Kant says in a reflection, “All philosophy has as its object reason: maxims, limits, and the end” (R 4987, 18:52). In this sense, again, the critical philosophy is, at its heart, merely the attempt to take the implicit *self-consciousness of reason* and transform it into something clearer and more explicit, which can serve as a foundation for further philosophical theorizing.

---

33 See (Boghossian 2012; 2014; Dogramaci 2017; Valaris 2017; 2019; Neta 2019; 2018). The reading of Kant as endorsing some form of a “taking condition” goes back at least to (Pippin 1987), compare (Allison 2004).

34 (Schapiro 2021).

35 I leave to the side here the important question of how this model applies to sensibility. For further discussion, see (Schafer forthcoming), which offers a sympathetic discussion of (Conant 2016; Boyle 2016).

36 Of course, this leaves unexplained why this provides us with a consciousness of the necessity of these principles. For discussion, see (Schafer 2019a).

37 Compare the post-Kantian methodology recently advocated by Sebastian Rödl, which he describes as follows:
Third, although these capacities do provide us with a consciousness of their own acts, principles, and ends, this consciousness does not rise to the level of genuine cognition (Erkenntnis) of these capacities (or the “soul” which underlies them). For instance, while this consciousness does provide us with a grasp of their fundamental principles of activity and ends, it is far too indeterminate to count as cognition of an object in the sense discussed in section 2. As Kant stresses in the discussion of the Paralogisms, it provides us with no basis for determining whether these capacities are capacities of a single spiritual substance or a network of cooperating substances or something else. As a result, as Kant sometimes puts the point, our consciousness of the acts and principles of these capacities remains merely “formal” as opposed to providing us with the “material” consciousness of real objects that is characteristic of cognition in a strict sense:

All rational cognition is either material and concerned with some object, or formal and occupied only with the form of the understanding and of reason itself and with the universal rules of thinking in general, without distinction of objects. (G 4:387)

Thus, we should not worry that this appeal to a consciousness of our basic rational capacities trespasses against Kant’s ban on theoretical cognition of things in themselves. For this consciousness is neither genuine cognition (in a “material” sense) nor, strictly speaking, theoretical in character. Rather, like all “merely formal” forms of self-consciousness, it involves a form of consciousness that is prior to the distinction between theoretical and practical cognition. For this latter distinction, insofar as it concerns the proper relationship between a cognition and its object, only arises within the domain of material cognition. In short,

This explains what may appear a curious character of the present essay: it propounds no theses, advances no hypotheses, does not recommend a view or position; it does not give arguments that are to support a view, it does not defend a position against competing ones, it does nothing to rule out contrary theses. It does nothing of the sort because it is—it brings to explicit consciousness—the self-consciousness of judgment. (Rödl 2018, 12–13)

On the most fundamental point, Rödl is in agreement with my reading of Kant – for both see foundational philosophy as bringing reason to full and explicit self-consciousness. Nonetheless, Rödl’s insistence that this task should not involve “giving arguments” or “defending positions against competing ones” conflicts with Kant’s view of reason as (in part) the faculty for mediate inference – that is for exactly the sorts of “arguments” that Rödl here disparages. To ignore this is to forget that the self-consciousness at issue here is the self-consciousness of reason.

Note the contrast between Kant and many of those influenced by him. Consider, for example, Fichte’s remark that, “Reason necessarily cognizes itself completely, and an analysis of its entire procedure, or a system of reason, is possible.” (GA I, 5:68-69) For Kant, to speak of reason’s “cognizing itself completely” would be to attribute powers to human reason that lie beyond it. For Fichte or Hegel (on the other hand) the idea that reason’s self-consciousness might be incomplete is incompatible with reason’s own demand for systematic comprehension.

As Matt Boyle notes, these limitations on “self-cognition” are closely related to Kant’s commitment to the principle that to know a power by its acts is to know it by its effects, rather than by its causes.

See (Engstrom 2016).
the sort of formal self-consciousness that is at issue here is not a form of material cognition – and so, a fortiori, it is not a form of theoretical (or practical) material cognition either.

For the critical Kant at least, this consciousness of the acts and principles of capacities is as deep as we can penetrate into the nature of things:

But all human insight is at an end as soon as we have arrived at basic powers or basic faculties; for there is nothing through which their possibility can be conceived, and yet it may not be invented and assumed at one’s discretion. (CPrR 5:46-7, my emphasis)

For Kant, our ability to develop a systematic picture of reality comes to an end with (our admittedly limited grasp) of the nature of these faculties. As a result, philosophy in its critical phase can arrive at the first principles that its second, doctrinal phase requires only by developing a systematic picture of our basic rational capacities. For it is the resulting system of principles, as the principles of our rational capacities, that provides philosophy with the unity that genuine comprehension requires, at least insofar as comprehension is possible for creatures with our cognitive limitations:

In this way the a priori principles of two faculties of the mind, the faculty of cognition and that of desire, would be found and determined as to the conditions, extent, and boundaries of their use, and a firm basis would thereby be laid for a scientific system of philosophy, both theoretical and practical. (CPrR 5:12, my emphasis; cf. CPJ 5:169)

The concepts of nature, which contain the ground for all theoretical cognition a priori, rested on the legislation of the understanding. ... The concept of freedom, which contains the ground for all sensibly unconditioned practical precepts a priori, rested on the legislation of reason. (CPJ 5:176)

In considering these passages, it is important to stress that what is fundamental for the critical Kant is not some single ur-faculty or ur-principle, but rather a system of interrelated faculties that cannot be reduced to any single faculty or principle. This, of course, sets Kant apart from the German Idealists who took the “completion” of Kant’s philosophical project to require just such an ur-philosophical-faculty-cum-principle. As such, it is important to stress that, at least for the critical Kant, the discovery of such an ur-principle is not required for philosophy to take on a rationally satisfying form. Rather, what is required is

41 Compare: “the chief question always remains, ‘What and how much understanding and reason cognize free of all experience?’ and not: ‘How is the faculty of thinking itself possible?’” (Axvi-xvii)
merely that these faculties form a rational system – that is, that their acts form a teleological unity organized under the overarching aims of reason as the highest faculty.\textsuperscript{42} So long as this is true, Kant claims that the principles these faculties provide to philosophy will have the unity that is required for philosophy to take on a rationally satisfying form.

5. The Philosopher as the Legislator of Reason

So, how exactly is philosophy to develop (i) a systematic picture of our rational capacities and, on that basis, (ii) a systematic account of the metaphysics of nature and of morals? The first step in this process, as we’ve noted above, involves reflection on the acts of these capacities – a reflection that aims to give us a clear and systematic grasp of their principles and aims. It is this sort of reflection that provides philosophy with the system of concepts it uses to make sense of both our rational capacities and, ultimately, nature and morality. For this reason, as Kant stresses, the role of definitions in philosophy is much more modest than it is in, say, mathematics.\textsuperscript{43} For, in philosophy, “definitions come about only as expositions of given concepts”. In other words, in philosophy, it is the self-conscious character of our basic rational capacities that provides us with the basic conceptual framework for philosophical thought – and the role of definition and conceptual analysis in that thought is simply to give us as precise and complete a grasp of these concepts as possible. Thus, unlike mathematical definitions which (in a sense) “make the concept itself” (through their relationship to construction in pure intuition), philosophical definitions “merely explain” the concepts which the self-conscious character of our rational capacities have already provided to us. (A730/B758)

This brings us to perhaps the most familiar question about Kant’s methodology – namely, the place of “transcendental arguments” in it.\textsuperscript{44} Such arguments are often thought to provide Kant’s philosophy with its foundations. So, for example, in his insightful discussion of Kant’s conception of our rational faculties, Haag recently writes that:

The existence of faculties, from the perspective of transcendental philosophy, likewise has to be established by reflecting on the conditions of the possibility of conscious experience. ... it is the function that transcendental reflection reveals as needing to be fulfilled that justifies the introduction of a particular faculty. (Haag 2015)

\textsuperscript{42} See (Schafer forthcoming). Cf. (Fugate 2014).
\textsuperscript{43} For an illuminating discussion of the role of definitions in philosophy see (Nunez 2019). For mathematical definitions in Kant, see (Dunlop 2012; Heis 2014; Sutherland 2010).
\textsuperscript{44} For an overview of such arguments, see (Stern 2020). For my take on their role, see (Schafer forthcoming).
In this passage, Haag provides us with an admirably clear articulation of a widespread view concerning Kant’s philosophical methodology. But if what I have been saying is correct, there is reason to be skeptical of certain elements in this understanding of the foundations of Kant’s system. In particular, while there is no doubt that Kant often justifies the introduction of faculties and capacities into his account on the basis that they are necessary conditions on the possibility of something else, there is reason to be skeptical that this is how Kant would justify his appeal to the most basic self-conscious rational capacities at work in his system – such as the understanding or reason.

To see this, it is helpful to remember that the starting point of Kant’s “transcendental reflections” in all three Critiques is not some minimal conception of “conscious experience” like what Haag refers to here – but rather something a good deal more robust – be this the sort of empirical cognition of objects that forms the focus of the first Critique or the sort of practical cognition of an unconditional moral law that forms the focus of the second. Thus, while the necessary conditions on these sorts of cognitive achievements are, of course, of great interest to Kant, the starting point of his investigations already attributes to us thoroughly non-trivial rational achievements and activities – and the capacity to perform them. As such, the attribution of these capacities to us cannot be based, for Kant at least, on a further “transcendental argument” – at least when such arguments are understood as justifying the introduction of some capacity (e.g. the capacity to X) on the basis of the fact that possession of this capacity is a necessary condition on some distinct Y. Rather, as suggested above, the attribution of these basic capacities to us is grounded merely in their character as self-conscious – and, in particular, in the consciousness they provide of their own characteristic forms of activity. As a result, in these most basic cases, if there is a sort of transcendental reflection at work here, it is simply the sort of reflection that is required to make clear and explicit the consciousness we already have of the activities and capacities that are constitutive of our form of self-conscious rationality.

Nowhere is this clearer than in Kant’s much maligned discussion of the “Fact of Reason” in the second Critique. Far from representing a regression to “a state of pre-critical dogmatism”, this passage actually presents us with an unusually clear instance of Kant’s critical methodology in action. After all, what we find there is precisely a practical philosophy that takes for granted only practical reason as a self-conscious rational capacity. That is, in this passage, it is the self-consciousness of pure practical reason that provides us with access to the moral law as the internal principle of that faculty’s activity.

45 Compare (Cassam 2016). This is, of course, not uncontroversial. For example, (Guyer 1987) attempts to read the Transcendental Deduction in a much more ambitious fashion, while (Korsgaard 1996) is often read as attempting something similar with respect to Kant’s practical philosophy.
The result, as noted above, is precisely, “a system that takes no foundation as given except reason itself” (Prol. 4:274) In this way, in the practical domain, Kant’s aim is to develop a systematic “metaphysics of morals” by treating practical reason’s principle as the fundamental principle that unifies and makes possible a genuine “science” of practical cognition. In the theoretical domain, Kant proceeds in a similar fashion in developing a systematic metaphysics of nature. For here too the principles that make this system possible are provided by the nature of our faculties for theoretical cognition. This is possible because of how transcendental idealism allows Kant to reconceive of the relationship between nature and cognition. The details of this are, of course, hotly debated. But the important point for present purposes is that here too we can see that Kant is operating within a “capacities-first” framework in developing a systematic account of the natural world. For here, much as in the practical case, Kant sees the possibility of such an account as dependent upon our rational capacities providing the philosopher with the non-trivial \textit{a priori} principles it requires. On the deepest level, it is this that drives Kant towards transcendental idealism – that is, to a view on which the fundamental “principles of being” are (on formal level) explained by the fundamental “principles of cognition”.

To say this, of course, is only to provide the barest outlines of how Kant develops the critical philosophy, starting from these capacities. But I want to close by discussing one other aspect of how Kant conceives of the critical method – namely, the connection between “capacities-first philosophy” in the sense sketched above and Kant’s Enlightenment ideal of reason as an \textit{autonomous} capacity.

The idea that \textit{only} reason (as autonomous) can sit in judgment of itself is, of course, fundamental to how Kant conceives of his philosophical project:

The critique … which derives all decisions from the ground-rules of its [reason’s] own constitution, whose authority no one can doubt, grants us the peace of a state of law … [In this] state of law it is the \textbf{verdict}, which, since it goes to the origin of the controversies themselves, must secure a perpetual peace. And the endless controversies of a merely dogmatic reason make it necessary to seek peace in some sort of critique of this reason itself, and in a legislation founded upon it... (A751-2/B779-80, compare Anth. 8:145)

\footnote{For an overview, see (Stang 2018).}
\footnote{Compare how (Kimhi 2018) situates Kant’s philosophy in terms of the relationship between “principles of being” and “principles of thinking”.}
Nothing here can escape us, because what reason brings forth entirely out of itself cannot be hidden, but is brought to light by reason itself as soon as reason’s common principle has been discovered. (Axx; cf. 28:1051)

As this makes clear, Kant does not share Hegel’s aim of constructing a truly presupposition-less philosophy. But this does not mean that philosophy can assume whatever it wishes for Kant. Rather, as Kant understands it, a truly critical philosophy can presuppose only the capacity of reason itself. In this sense, the critical philosophy is simply reason’s self-critique. Or rather, since reason’s self-critique began long before Kant, it is this self-critique in a genuinely self-conscious and systematic form.
References

All references to Kant’s works use the standard academy pagination with abbreviations of the title, with the exception of references to the first Critique, which use the standard A/B pagination.
