

# 1 Kant's Conceptions of Reason and Metaphysics

---

In this chapter, I will offer a first and preliminary outline of Kant's accounts of reason and metaphysics and introduce various themes that will be developed further in later chapters.<sup>1</sup>

## 1.1 Kant's Conception of Reason

Reason, for Kant, is primarily a mental capacity, or 'faculty of the soul' (*Seelenvermögen*) (5:177) – that is, an ability of thinking beings to be in mental states of different kinds.<sup>2</sup> More specifically, reason is a *cognitive* faculty (*Erkenntnisvermögen*) that allows one to have 'objective representations': mental states aimed either at truly representing things or at actively bringing things about (in the widest possible sense of 'things'). Human beings have various cognitive capacities, distinguished by the kinds of representations to which they give rise (or, more generally, by the cognitive functions they perform): sensibility, understanding, imagination, power of judgment, reason (and more). Kant has different ways of classifying these capacities depending on which of their features he is interested in.<sup>3</sup> With respect to reason, Kant employs two such classifications: on the one hand, cognitive capacities can be classified as resulting either in a priori or in empirical cognition (e.g. A835–6/B863–4); on the other hand, they can be distinguished according to whether they represent things intuitively (in human beings this is the case with sensible

<sup>1</sup> Note that my aim in this chapter is not to present an original interpretation, but merely to lay the ground for what follows.

<sup>2</sup> Even though Kant sometimes uses the terms '*Fähigkeit*' (capacity), '*Vermögen*' (faculty) and '*Kraft*' (power) interchangeably, in other places he draws subtle but important distinctions between them. For instance, while 'faculty' (*Vermögen*) involves an activity of the subject, 'capacity' does not (Heßbrüggen-Walter 2004). I will return to the specific connotations of the term '*Vermögen*' in Kant in Section 1.1.2.

<sup>3</sup> In the Introduction to the *Critique of the Power of Judgment*, Kant distinguishes three basic mental faculties: the faculty of cognition, the faculty of desire, and the feeling of pleasure and displeasure (5:177). While Kant sometimes treats *practical* reason as part of the faculty of desire (*Begehungsvermögen*) (5:24), at other times he regards it as a *cognitive* faculty (*Erkenntnisvermögen*) (5:174), namely the faculty of *practical* cognition (cognition of what one ought to do).

representations) or discursively (in intellectual representations or concepts) (e.g. A68/B93). (As a first approximation, we can say that intuitive representations represent particulars, while discursive representations represent general features of things; 9:91; 8:399.)

Now *reason*, in the sense most relevant to Kant's account of metaphysics, is a capacity of a *a priori discursive* cognition. It shares this characterization with the pure understanding (the capacity for a priori conceptual thought and judgment), from which it is distinguished by its characteristic type of activity (and its corresponding type of representation), namely the drawing of mediate (indirect) *inferences*.<sup>4</sup>

### 1.1.1 Reason as the Capacity for Mediate Inference

Reason, in this sense, is the capacity to logically derive particular cognitions from more general ones by means of intermediary cognitions, as in the inference: 'All humans are mortal; all philosophers are human; thus, all philosophers are mortal.' This inference is based on the schema 'All A are B; all C are A; thus, all C are B,' the validity of which can be grasped a priori. At the same time, all its constituent representations (A, B, C) are concepts and as such are general, not singular. Thus *reason*, in the most fundamental sense, is the capacity for a *a priori discursive cognition by means of mediate inference*. It can hardly be denied that human beings possess such a capacity, even though, of course, we are fallible in its employment.

While Kant defines reason in terms of apriority, discursivity, and mediate inference (e.g. A298–301/B355–8), this conception of reason must be distinguished from two wider conceptions that Kant also employs without always clearly marking the difference. First, Kant sometimes uses the terms 'reason' (*Vernunft*) and 'understanding' (*Verstand*) interchangeably for the capacity of discursive cognition in general.<sup>5</sup> Second, Kant sometimes defines reason as the capacity for a *a priori* cognition (A11/B24) or as the 'entire higher faculty of cognition' (A835/B863) – a characterization that includes not just a priori

<sup>4</sup> Guyer and Wood translate '*Vernunftschluss*' as 'syllogism.' Since '*Vernunftschluss*' also includes hypothetical and disjunctive inferences (9:121–2), which are not syllogisms narrowly conceived (as instances of the four syllogistic figures, see Bennett 1974: 259), and in order to emphasize the essential role of reason, I will use 'rational inference' and, depending on the context, 'inference of reason' as translations of Kant's term '*Vernunftschluss*.'

<sup>5</sup> Compare, for instance, Kant's distinction between "the two stems" of human cognition, sensibility and understanding (A15/B29), with his claim that one of these "two stems" is "reason" (A835/B863). Already in his pre-critical writings, Kant distinguishes between reason and understanding without regarding them as distinct cognitive faculties. In *False Subtlety*, he claims that "*understanding* and *reason*, that is to say, the faculty of cognizing distinctly and the faculty of drawing rational inferences, are not different *fundamental faculties*. Both consist in the capacity to judge; but when one judges mediately, one draws an inference" (2:59).

cognitions of the 'pure' understanding (such as the principle that every alteration has a cause), but even those of mathematics (which, according to Kant, are not discursive but intuitive). Thus, Kant sometimes understands reason solely in terms of discursivity, at other times solely in terms of apriority.<sup>6</sup> What is at stake in Kant's Rational Sources Account, however, is the narrower conception of reason in terms of discursivity *and* apriority (*and* mediate inference).

As we will see later, Kant goes on to enrich this characterization of reason by showing that each of its three features includes further characteristics: discursivity means that any cognition based on reason can relate to its object only indirectly (A320/B377); apriority, for Kant, implies necessity and universality (B3); and the logical use of reason in rational inferences gives rise to a priori concepts of 'totalities,' that is, of all the objects of a certain domain (A321/B378). Moreover, the structure of rational inferences is iterative: the conclusion of one inference can be made the premise of another, and so on (A307/B364). By making use of this feature, reason, according to Kant, aims at placing *all* our cognitions in *one* inferential network, thus transforming them into a unified system the structure of which can be determined a priori (A302/B359; A832/B860). But before we turn to this fuller conception of reason (which is already part of Kant's Rational Sources Account), we have to consider what is, and what is not, involved in thinking of reason as a *faculty*.

### 1.1.2 Reason as a Faculty (*Vermögen*) and Reason as a System of Principles

Kant's wording can sometimes suggest that he thinks of reason, the understanding, and other 'faculties' as so many agents, each pursuing its own aims and acting in specific ways to realize them. For instance, the understanding is said to use concepts to form judgments (A68/B93); the imagination is said to synthesize a manifold of sensible representations (A78/B103). Similarly, reason is said to aim at bringing systematic unity to the manifold of empirical cognitions (A302/B359). There is something misleading about this way of speaking, but it also contains an important insight. What is misleading is that it can suggest that the cognitive faculties really *are* individual agents – homunculi that are active (in contrast to the human beings whose capacities they are). But of course this is not what Kant means, and nothing in what he says commits him to such a picture. Rather, 'faculties' (*Vermögen*) are always the faculties of a thing or a substance – in this case, the cognitive faculties of

<sup>6</sup> On the difference between these conceptions of reason, see Willaschek 2013.

human beings.<sup>7</sup> Thus when Kant says, for instance, that reason tries to bring the greatest manifold of empirical cognitions to the smallest number of principles (A305/B361), he does not want to deny that it is human beings, in virtue of having reason, who ‘bring cognitions to principles.’ **Speaking of faculties as if they were agents can primarily be understood as a literary device meant to facilitate exposition.**<sup>8</sup>

But economy of expression is not the only reason that leads Kant to speak of faculties as if they were agents. **Another reason, which embodies the insight connected with the faculty idiom, is that this allows him to describe cognitive activity in normative and teleological terms.** Rational human beings do not just *happen* to bring cognitions to principles; rather, this is what they *ought* to do. In thinking of cognitive performances as employments of cognitive faculties, Kant attributes a teleological structure to them that combines descriptive and normative elements. The descriptive element is highly complex, consisting in a variety of more specific cognitive dispositions to think in ways that can be characterized as rational. For instance, rational beings tend to draw inferences based on *modus ponens*, *modus tollens*, etc. The normative element is complex too, involving, on the one hand, standards for the correct employment of reason, such as inference rules. (While it is *correct* to infer ‘B’ from ‘If A, then B’ and ‘A,’ it is *incorrect* to infer ‘A’ from ‘If A, then B’ and ‘B.’) On the other hand, according to Kant, reason brings with it its own ends (e.g. A839/B867), needs (e.g. A309/B365), and interests (e.g. A462/B490–A476/B504). Thus, reason *aims* at systematic unity (we will soon turn to why this is the case) and takes an ‘interest’ in the truth of certain theses. **While it may be unclear whether Kant wants to commit himself to an Aristotelian picture of faculties as *dynamēis* that are teleologically geared toward their own realization (*energeia*), it is important to understand that in attributing reason to a human being, according to Kant, we not only attribute certain dispositions of thought to her but also hold her to certain standards, attributing specific ends and interests to her – or, more correctly, we normatively *require* her to have them.**<sup>9</sup>

Of course, this is a requirement of a rather weak and conditional kind. Thus, Kant does not hold that *every* rational being is *always* required to follow *all* rational principles and pursue *all* of reason’s ends. There are ends and principles of reason that are obligatory for everyone and anytime, namely moral ones. But when it comes to non-moral ends, specifically to theoretical ends

<sup>7</sup> See Heßbrüggen-Walter 2015. On the metaphysical implications of Kant’s account of cognitive faculties, see Heidemann 2017.

<sup>8</sup> For a recent defense of the use of this expository device in psychology, see Kahneman 2011.

<sup>9</sup> For ‘Aristotelian’ readings of Kant’s account of the cognitive faculties, see Kern 2006; Engstrom 2009. On the ‘conative’ character of reason, see Kleingeld 1998a. For a reading of Kant’s conception of reason that emphasizes its teleological features, see Ferrarin 2015.

(such as bringing one's cognitions under ever higher principles), this is obviously something one is only required to do under certain conditions.<sup>10</sup> *If* one reflects on one's empirical knowledge, and *if* one is free to follow through with one's reflection (to 'persevere' with it, as Kant puts it at A584/B612), *then* as a rational being one ought to care about the inbuilt *telos* of this kind of rational reflection. Of course, the plausibility of this claim will depend on the ends attributed to reason. Avoiding inconsistencies and contradictions may not be among the most important things in the world, but in not caring about them at all a rational being may justly be said to make a mistake that he or she *ought* to avoid.<sup>11</sup> Whether this also holds for the systematic unity of cognition and other ends that Kant attributes to reason is a question to which we will return (Chapters 2 and 5). In any case, Kant is committed to such a normative (and teleological) conception of reason, and for him it is part and parcel of thinking of reason not just as a disposition but as a faculty.

A consequence of this normative conception of reason is that reason is not just a psychological capacity but can also be viewed as a *body of cognitions and principles*. Kant speaks of reason 'considered subjectively as a human faculty of cognition' (A297/B353), thereby implying that reason can also be considered 'objectively,' that is, with respect to the principles and cognitions it contains – as rules that implicitly guide our thinking and as conclusions we can arrive at solely on the basis of a priori reasoning. No matter what rules of thinking and reasoning a rational being in fact follows, there are those it *ought* to follow. And no matter which conclusion it in fact arrives at, there are (so Kant claims) those it *ought* to arrive at (*if* it follows certain rational trains of thought to their end). Moreover, no matter what ends a rational being in fact pursues, there are those, qua rational, that it *ought* to pursue. Thus, considered objectively, reason is a system of a priori principles and cognitions, as well as ends, needs, and interests (which, when restricted to *pure* reason, Kant calls the "system of pure reason"; A841/B869). Reason in this sense is not distinct from reason as a faculty but rather a different perspective on the same mental faculty: while reason can be considered 'subjectively' as a capacity to think in accordance with certain principles in order to pursue certain ends and thereby to arrive at certain conclusions, it can also be considered 'objectively,' with respect to the principles and ends that guide our use of that capacity and the conclusions to which its employment leads.

<sup>10</sup> Kant repeatedly contrasts the conditional character of speculative ends with the unconditional character of moral ends (e.g. 5:5; 5:142; 8:139; 9:87).

<sup>11</sup> But see Kolodny 2005, who argues that there is no normative requirement to care about consistency. It seems that if there is such a requirement, it must derive from an interest in truth, since inconsistency among theoretical attitudes tells us that at least one of them must be false. Kant takes consistency to be constitutive of reason (5:120).

### 1.1.3 *Speculative and Practical Reason, Pure Reason, and Universal Human Reason*

Within his conception of reason as a faculty of a priori discursive cognition based on rational inferences, Kant draws two important distinctions: between *theoretical* (and more specifically, *speculative*) and *practical* reason on the one hand, and between reason *in general* and *pure* reason on the other.

Unsurprisingly, while theoretical reason is concerned with theoretical questions (questions about ‘what is’ the case; 9:86), practical reason is concerned with practical questions (about ‘what ought to be’; *ibid.*). Theoretical reason uses rational inferences to derive theoretical conclusions from theoretical principles, while practical reason derives practical propositions (and ‘actions’; 4:412) from practical principles (typically in conjunction with theoretical premises). Instead of ‘theoretical reason,’ Kant sometimes speaks of ‘speculative reason.’ Although he does not always seem to distinguish between the two, he uses the term ‘speculation’ mainly for a specific application of theoretical reason, namely that which results in cognitions that are ‘abstract’ (not ‘concrete’) (4:369; 9:27), lack practical applications (9:86), and go beyond the limits of possible experience (A634–5/B662–3).<sup>12</sup> Speculative reason in this sense contrasts with “universal human reason” (B22) or “common human reason” (Aviii; B424)<sup>13</sup> (*allgemeine* or *gemeine Menschenvernunft*), which needs concrete “images” (9:27), is interested in practical consequences, and generally stays within the limits of experience. While all human beings make use of reason, only philosophers (both professional and amateur) engage *speculative* reason. As Kant emphasizes (e.g. Avii; 9:27), however, common human reason and speculative reason are continuous insofar as some of the questions ordinary people ask themselves when thinking about concrete empirical issues of practical relevance (e.g. about what caused the fire that burned down a house) can lead quite naturally to speculative questions (e.g. whether there are uncaused causes). Indeed, this continuity is an essential aspect of the Rational Sources Account, which, as we have seen, is the claim that metaphysical questions “arise from the nature of universal human reason” (B22), which means that they are not the arbitrary inventions of philosophers but rather grow naturally out of ordinary ways of thinking.

The other important distinction Kant draws within the conception of reason is the one between reason in general and *pure* reason in particular. As we have seen, reason itself is the capacity of a priori cognition, which implies that its

speculative use  
of reason

universal  
human reason

pure reason

<sup>12</sup> On these three ways of distinguishing between ‘theoretical’ and ‘speculative,’ see Lau 2015.

<sup>13</sup> Guyer and Wood translate “*gemeine Menschenvernunft*” in the A-Preface (Aviii) as “ordinary common sense,” thereby obscuring the fact that Kant means the same ‘human reason’ that is the topic of the first paragraphs of the A-Preface (and of the book as a whole).

own guiding principles are a priori too. Nevertheless, people mostly apply their reason to *empirical* questions, for instance by deriving empirical conclusions from empirical premises. *Pure* reason, by contrast, is concerned exclusively with deriving a priori cognitions from a priori principles. This general distinction can be spelled out in different ways, however. Thus, in the Introduction to the *Critique of Pure Reason*, Kant first defines reason as “the faculty that provides the principles of cognition a priori” and then adds: “Hence pure reason is that which contains the principles for cognizing something absolutely a priori” (A11/B25). Here, Kant uses the term ‘reason’ in a wide sense which also encompasses pure intuition and the pure understanding, since both provide us with principles of cognition a priori. But if that is what reason is, what does Kant mean by “*pure* reason”? It seems plausible that he intends his distinction between reason in general and pure reason to parallel the distinction, a little earlier in the text, between a priori cognitions and *purely* a priori cognitions (B3).<sup>14</sup> While a priori cognitions are judgments whose ‘objective validity’ we can come to recognize a priori, independently of experience, cognitions that are *purely* a priori are those whose constitutive representations are also a priori (that is, not derived from experience). Kant’s example is the judgment “Every alteration has its cause” (B3), which we can cognize a priori but which is not *purely* a priori because the concept of alteration is empirical.<sup>15</sup> This would mean that pure reason is the faculty of *purely* a priori cognition – of cognition that is independent of experience in terms of how we acquire its constituent (sub-judgmental) representations and how we can come to recognize their ‘objective validity.’ Pure reason in this sense is distinguished from reason in general by its more limited scope (*purely* a priori cognitions instead of a priori cognitions in general).

Compare this with the way Kant distinguishes between reason in general and pure reason in the Introduction to the *Transcendental Dialectic*, in a section entitled “On the Pure Use of Reason”:

Can we isolate reason, and is it then a genuine source of concepts and judgements that arise solely from it and thereby refer it to objects . . .? In a word, the question is: Does reason in itself, i.e. *pure reason*, contain a priori synthetic principles and rules, and in what might such principles consist? (A305–6/B362–3; emphasis added)

As Kant explains, these principles would have to differ from the a priori principles of the understanding in being not only “cognition from concepts” – that is, discursive – but rather cognition “from *mere* concepts” (A307/B364,

<sup>14</sup> Parts of this and the next paragraph are adapted from Willaschek 2013; see there for a more detailed interpretation of Kant’s two accounts of ‘pure reason.’

<sup>15</sup> Unfortunately, Kant says a little later (B5) that this judgment is purely a priori, but we can ignore this here.

emphasis added; see also A301/B357; A302/B258) – that is, independent even of the a priori forms of intuition (space and time). It is in this independence not only from experience but from anything belonging to sensibility and intuition that the purity of reason consists. So while pure reason, according to the Introduction to the *Critique of Pure Reason*, is the faculty of purely a priori cognition, the Introduction to the Transcendental Dialectic defines it even more narrowly as the faculty of a purely *discursive* (i.e. conceptual) a priori cognition. As we will see, it is this narrower conception of pure reason that is most relevant to the Rational Sources Account.<sup>16</sup>

Let me close this section with a remark on Kant's 'multiple' conceptions of reason. Even though Kant often speaks of 'speculative,' 'practical,' 'pure' *reason*, thereby suggesting that these are distinct faculties (or perhaps sub-faculties), Kant just as often speaks of 'speculative,' 'practical,' etc. *uses* of reason, and it seems that he took this to be the philosophically more adequate way of speaking (of which the other formulation is a mere abbreviation). Thus when Kant defines pure reason in the Introduction to the Transcendental Dialectic, the section title reads 'On the Pure Use of Reason' (A305/B362), and he says in the *Critique of Practical Reason* that pure reason, in turn, can be regarded "in its speculative or practical use" (5:107).<sup>17</sup> This shows that Kant thinks of pure, speculative, and practical reason not as distinct cognitive faculties, but rather as different employments of the same faculty, namely the faculty of arriving at cognitions through rational inferences. As Kant emphasizes in the *Critique of Practical Reason*, "it is still only one and the same reason which, whether from a theoretical or a practical perspective, judges according to a priori principles" (5:121).<sup>18</sup> So the overall picture is that human beings have a capacity for logical reasoning that can be applied to different subject matters (including questions about what to do and how to act). Possession of this capacity brings with it certain normative commitments (concerning consistency, systematicity, etc.). Moreover, in its 'pure' aspect it

<sup>16</sup> Both ways of distinguishing between reason in general and pure reason offered by Kant in the first *Critique* (1781) seem to be meant to apply only to theoretical or speculative reason, not to practical reason. This should not surprise us since Kant first mentions *pure practical* reason in the *Groundwork* (1785) and fully develops his account of pure practical reason only in the second *Critique* (1788). I discuss Kant's conception of practical reason and his distinction between pure practical reason and empirically conditioned practical reason in Willaschek 1992 and Willaschek 2006.

<sup>17</sup> Thus, 'pure practical reason' is short for 'the pure and practical use of reason.'

<sup>18</sup> It is an open question how pure practical reason and speculative reason, according to Kant, *can* be employments of the same faculty, given that they seem to work in accordance with very different principles and to have different ends. On the problem of the so-called unity of reason, see e.g. Kleingeld 1998b; Timmermann 2009. Note that this 'unity of reason' (*Einheit der Vernunft*) (the unity among the different employments of reason) is different from the 'unity of reason' (*Vernunftseinheit*) that reason is supposed to bring into the manifold of our cognitions (A309/B365).



can be used to derive a priori conclusions from a priori principles, which raises the question whether these principles and conclusions, their lack of sensible content notwithstanding, can be regarded as a priori *cognitions* (assuming, with Kant, that cognitions must relate to objects and that only sensible content bears a direct relation to objects; see Section 9.3).

#### 1.1.4 Two Aspects of Reason: A Historical Digression

Kant's conception of reason is part of a long and highly complex tradition. Before we proceed to Kant's distinction between the logical and the real use of reason, it may be helpful to take a cursory glance at its historical background.

What has traditionally been called reason, or rational thought, comprises two clearly distinguishable aspects whose relation has long been a matter of controversy (Horn and Rapp 2001). On the one hand, there is the intuitive grasp of abstract or general truths (truths that cannot be apprehended by the senses); on the other, there is discursive reasoning (the logical progression from premises to conclusions). Thus, Plato distinguishes between *noësis* (rational insight into the ideas) and *dianoia* (discursive reasoning), which he seems to have understood as applications of the same faculty, namely the faculty of reason or *logos* (e.g. *Politeia* 511b–e). Building on Plato's distinction, Aristotle distinguishes between a capacity for rational insight into first principles, which he calls *nous*, and a capacity for deductive knowledge (*epistēmē*), which he at least sometimes seems to subsume under *logos* (e.g. *Nicomachean Ethics* 1143a36–1143b1).<sup>19</sup> Later philosophers, writing in Latin, translated *noësis* as 'intellectus' and *dianoia* as 'ratio,' although terminology in this field is varied and inconsistent (Horn and Rapp 2001). Many early modern authors, particularly of the rationalist tradition, distinguish between insight into principles and logical reasoning as two distinct mental *activities*, attributing these activities not to two different *faculties* but rather to one faculty for which the terms 'intellectus,' 'ratio,' and 'ingenium' are often used interchangeably (Horn and Rapp 2001). As Descartes puts it, there are only two ways of gaining knowledge and certainty through the intellect (*intellectus*), namely intuition (*intuitus*) and deduction (*deductio*) (*Regulae*, Rule 3, §4).

In Germany, starting with the German works of Meister Eckhart, the terms 'Verstand' and 'Vernunft' were used to translate 'intellectus' and 'ratio,' with some authors translating 'intellectus' as 'Vernunft' and 'ratio' as 'Verstand,' and others adopting the converse convention (Horn and Rapp 2001). Christian Wolff, for instance, follows Leibniz in defining 'Vernunft' (reason) as "the faculty of seeing into the connection of truths" (*Deutsche Metaphysik*, §368;

<sup>19</sup> While the assumption that *noësis* is intuitive has been disputed (e.g. Horn and Rapp 2005), it clearly represents the traditional and historically most influential reading.

Watkins 2009: 30), whereas ‘*Verstand*’ (understanding) is more generally defined as “the faculty of distinctly cognizing what is possible” (*Deutsche Metaphysik*, §277; Watkins 2009: 24). For Wolff, as for his critic Crusius (*Entwurf*, §441; Watkins 2009: 176), reason and understanding are not two distinct faculties; rather, reason is a special application of the more general and encompassing faculty of understanding.

Setting the historical and terminological complications aside, we can detect a fairly consistent distinction between two aspects of a priori reasoning. **On the one hand, there is (what from a present-day perspective can be characterized as) the capacity for the truth-preserving progression from a given set of propositions to other propositions not included in that set. This capacity is concerned not with the truth of single propositions, but rather with the necessary relations between the truth of one or more propositions and the truth of others.** In this sense, the knowledge conveyed by this kind of logical reasoning is always *conditional*: given the truth of some propositions, the truth of other propositions follows. **On the other hand, there is the capacity to grasp the truth of a principle intuitively. This kind of rational insight is not conditional in the same way logical reasoning is. Rather, it is directed at one proposition at a time.** It is the capacity to know whether a proposition is true simply by understanding it, by grasping its content. Descartes, like other philosophers before him, compares this way of coming to know the truth of a proposition to vision (the ‘*natural light*,’ see e.g. the Third Meditation). Like seeing with one’s eyes, this purely mental, non-sensible seeing is not discursive (step-by-step, mediated through other cognitions, made up of elements that are available prior to it) but intuitive (instantaneous, immediate, holistic). And it does to the mind of the philosopher just what ordinary seeing (which they say ‘is believing’) does to the mind of the ordinary person: it commands assent. The paradigm of this kind of rational insight is grasping the truth of mathematical axioms.

Meier

We find this distinction at work in a text Kant must have known by heart, namely Meier’s *Auszug aus der Vernunftlehre* (Extract from the Doctrine of Reason), on which Kant based his lectures on logic from 1757 until his retirement from teaching almost forty years later in 1796. In §116, Meier defines reason (*Vernunft*) as the “faculty of distinct insight into the nexus of things.”<sup>20</sup> Later, in §318, he writes:

In a demonstration from reason all grounds the proof is based on<sup>21</sup> must be completely certain (§§193, 204); hence they are either demonstrable or indemonstrable (§313). In

<sup>20</sup> “[E]in Vermögen . . . , den Zusammenhang der Dinge deutlich einzusehen” (§116, 16:30); translations are my own.

<sup>21</sup> “[A]lle Beweisthümer,” which according to *Grimmsches Wörterbuch* is synonymous with ‘Beweisgrund’ (on which term see Chignell 2009).

the first case, they in turn must be proven. Consequently, a proof will not become a demonstration [from reason] until I arrive at indemonstrable grounds only. Empty [i.e. tautological] judgments, the fundamental judgments [i.e. axioms] and the postulates, are therefore the first starting points of all demonstrations from reason (§§314; 315). When the proof has been pursued up to judgments of this kind, the understanding finds complete rest.<sup>22</sup>

Thus, Meier distinguishes between derived and underived truths, that is, between truths we recognize on the basis of discursive reasoning and those we grasp immediately. He insists that 'demonstrations from reason' must ultimately rest on the latter. Underived, indemonstrable judgments are such that "their truth becomes clear from themselves, as soon as we have cognized them distinctly" (§313). And while 'reason' is the faculty of logical inference that provides the proofs, it is the 'understanding' that finds rest in indemonstrable judgments.

There are three things I would like to take from this brief historical digression. First, there is a long tradition of distinguishing between two different aspects of reason, namely rational insight into principles and logical reasoning. What both have in common, minimally, is that they are (real or merely purported) sources of non-empirical knowledge. Second, it is controversial whether these two sources of knowledge are fundamentally of the same type, and are thus applications of the same cognitive faculty, or whether they are fundamentally distinct, in that rational insight is intuitive whereas logical reasoning is discursive. Third, while German philosophers of the eighteenth century such as Wolff and Baumgarten use '*Vernunft*' as a name for the faculty of logical reasoning or of cognizing the 'concatenation' of truths, they think of it as an expression of the fundamental faculty of understanding, which also provides rational insight into general truths.

### 1.1.5 *The Logical and the Real Use of Reason*

It is only after some 300 pages that Kant, in a book that is, after all, entitled *Critique of Pure Reason*, explains in some detail what he means by 'reason.' And Kant admits: "Since I am now to give a definition of this supreme faculty of cognition [i.e. of reason], I find myself in some embarrassment" (A299/B355).

<sup>22</sup> "In einer Demonstration aus der Vernunft müssen, alle Beweisthümer, völlig gewiss sein §. 193. 204; sie sind also entweder erweislich oder nicht §. 313. In dem ersten Falle müssen sie wieder bewiesen werden. Folglich wird ein Beweis nicht eher eine Demonstration, bis ich nicht auf lauter unerweisliche Beweisthümer komme. Die leeren Urtheile, die Grundurtheile und Heischeurtheile sind demnach die ersten Anfänge aller Demonstrationen aus der Vernunft §. 314. 315. Alsdenn beruhiget sich der Verstand völlig, wenn der Beweis bis auf solche Urtheile fortgeführt worden" (§318, 16:91). In his copy of Meier's book, Kant wrote in the margin next to this paragraph the words "mathematical method" (Ref. 3124, 16:670), thereby linking the paragraph to the Cartesian method of deriving all truths from axioms which in turn are certain without demonstration.

The problem is that reason has both a merely formal or logical use, which abstracts from all content, and a real or transcendental use, which is “the origin of certain concepts and principles, which it [reason] derives neither from the senses nor from the understanding” (ibid.). The first use, Kant continues, has “long since been defined by the logicians as [the faculty] of drawing inferences mediately . . . ; but from this we get no insight into the second faculty, which itself generates concepts” (A299/B355).

What Kant has in mind here is clearly some version of the traditional distinction between *dianoia* and *noësis*: reason, on the one hand, as the capacity for logical reasoning or, more specifically, mediate inference, and reason, on the other hand, as the capacity for insight into non-empirical principles. To be sure, Kant does not attribute to reason the ability to intuit the truth of non-empirical principles that has traditionally been associated with *noësis*, *nous*, or *intellectus*; rather, he speaks of reason as containing “the origin of certain concepts and principles” (A299/B355). But this is merely because Kant does not want to commit himself to the view that reason is actually *successful* in its attempt to gain purely rational insight into first principles. By saying that reason is, or contains, “the origin of certain concepts and principles,” what he wants to say is that reason, in its real or transcendental use, is at least a *purported* source of (substantial, not just formal) non-empirical knowledge. It is this ‘real’ use of reason which is at stake in the Rational Sources Account.<sup>23</sup>

Whereas the distinction between the logical and the real use of reason echoes the Platonic distinction between *dianoia* and *noësis*, Kant regards both as expressions of a single faculty of reason, thereby placing himself more specifically in a *Cartesian* tradition; as we have seen, for Descartes (and many of his rationalist followers) the one faculty of reason or intellect comprises two distinct applications that structurally parallel Kant’s logical and real use of reason.

But if reason has these two very different employments, how can it be characterized such that we can understand how these employments are

<sup>23</sup> In contrasting the logical and the real use of reason and their respective principles, Kant often uses the word ‘transcendental’ rather than ‘real’ (e.g. A299/B356; A306/B363), thus treating the terms ‘logical’ and ‘transcendental’ as antonyms. This can be confusing in light of the fact that Kant also distinguishes between general and transcendental logic (A55/B79–80), in which case ‘logical’ and ‘transcendental’ are not antonyms. But there is an obvious parallel that explains Kant’s usage: while general logic abstracts from the objects of cognition, transcendental logic does not (A55/B79–80); similarly, while the logical use of reason abstracts from objects, the real or transcendental use does not (A299/B355). Thus, when it is contrasted with ‘real’ or ‘transcendental,’ Kant uses ‘logical’ as pertaining to ‘general logic.’ On the other hand, when Kant speaks of transcendental as opposed to logical principles, he does not necessarily mean that they are ‘conditions of the possibility of experience’ or part of an explanation of a priori cognition (B25), but rather that they concern objects (see e.g. Caimi 1995: 309; Guyer 2003: 278).

nevertheless two uses of the same faculty? Kant suggests a definition of reason that is meant to cover both its logical and its real use, namely reason as the *faculty of principles*, where 'principle' can mean either the general premise or *major* of a syllogism ("comparative principles") or "synthetic cognitions from concepts" ("principles absolutely" so called) (A301/B357–8). The latter would be a priori principles based on reason alone. Now this may seem to be a merely verbal maneuver since the two kinds of 'principle' are clearly very different. It is only in what follows that Kant explains how he thinks of the logical and the real use of reason as uses of the same faculty (and how comparative and absolute principles are related).<sup>24</sup> His general idea, which we will explore in detail in later chapters, is that the real use of reason grows naturally out of its logical use and that we inevitably move from using comparative principles to assuming absolute ones.

the broad or inclusive definition of reason

While the *logical* use of reason (more on which in Chapter 2) abstracts from content and is concerned with the logical entailment relations between judgments, the *real* use of reason (more on which in Chapter 3) "aims at objects" (*auf Gegenstände geht*) (A306/B363). That is, it goes beyond the logical relations between judgments and aims at cognizing objects (in the widest sense of the term). Thus, 'real' (from Latin *res*, thing) here means 'object-related' or 'concerning not just representations but things.'<sup>25</sup> However, both uses of reason share a concern with bringing unity to the manifold cognitions provided by the senses and the understanding: the logical use of reason by searching for its 'principles' (that is, more and more general premises) from which cognitions supplied by the understanding can be derived, the real use of reason by looking for its 'principles' (that is, for fundamental aspects of reality) that ground, or explain, what is less fundamental (conditioned, dependent).<sup>26</sup>

<sup>24</sup> Kant does not align the logical use of reason with comparative principles and the real use with principles in the latter 'absolute' sense, thereby allowing for 'comparative' and 'absolute' principles in both the logical and the real use of reason.

<sup>25</sup> Kant's distinction between 'logical' and 'real,' where the former is formal and concerns cognitions in abstraction from their objects and the latter is material and concerns objects, goes back at least to the 1755 essay *Nova Dilucidatio*, where Kant distinguishes between "logical opposition" (contradiction) and "real opposition" (real repugnance) (2:171) and between "logical ground" and "real ground" (2:202). It reappears for instance in Kant's inaugural dissertation, now in the form of a distinction between the "logical use" and the "real use" of the understanding (2:393), which is a direct predecessor of the distinction between the logical and the real use of reason. A late version of that distinction in slightly different terminology can be found in the essay *On a Discovery*, where Kant distinguishes between a "logical (formal)" and a "transcendental (material)" reading of the Principle of Sufficient Reason (8:193), where the original opposition "logical/real" is also used, but only with respect to reasons (8:198).

<sup>26</sup> The term 'principle,' which is the translation of the Greek *'archê'*, can refer to both fundamental premises and reasons (on the 'logical' side) and first causes and grounds (on the 'real' side). See Hebbeler 2012 for a helpful account of Kant's conception of principles in the first *Critique*.

Even though the real use of reason ‘aims at objects,’ it is meant to work completely a priori and discursively, independently of anything sensible. The a priori principles of the *understanding* depend on something given in sensible intuition insofar as they, according to the Transcendental Analytic, are valid only for objects of possible experience (A158/B197). By contrast, cognition resulting from the real use of *reason* would have to be cognition ‘from mere concepts’ (that is, completely discursive, independent of even the a priori forms of intuition: space and time; A307/B364). Because of this radical independence from intuition and sensibility, Kant can identify the ‘real’ use of reason with its ‘pure’ use (the section title at A305/B362).<sup>27</sup>

At the same time, Kant emphasizes the *synthetic* character of the principles and cognitions in which the real use of reason, if successful, would result. After all, analytic cognitions do not tell us anything specifically about objects, but only about the content of our concepts. If the real use of reason is to consist in gaining cognition and knowledge of *objects* (not just, as with the logical use, of inferential relations between our cognitions), it must result in synthetic principles and cognitions (B18).<sup>28</sup> However, in the Transcendental Analytic Kant had argued that synthetic cognition cannot be purely discursive, but always requires some intuitive element (minimally, a relation to *possible* experience). Reason, by contrast, is a purely discursive faculty for Kant, which means that rational insight into first principles cannot be based on anything intuitive (which in human beings is always sensible), but only on logical reasoning and the discursive principles and concepts that come with it. Already here, at the very beginning of Kant’s investigation into the real use of reason in the Transcendental Dialectic, we can therefore foresee that this story will not end well: while according to the Transcendental Analytic there cannot be cognition from concepts alone, according to the Dialectic the cognitions of pure reason would have to be precisely that: purely discursive, cognitions from mere concepts.

In distinguishing between the logical and the real use of reason, Kant does not want to claim that they are unrelated. Rather, the real use builds on the logical use insofar as its concepts and principles correspond to, and perhaps can even be derived from, the forms and principles that characterize the logical use of reason. Conversely, the logical use of reason, when considered in isolation (as we do in formal logic), can be regarded as a mere abstraction from the way we reason about specific objects and events in science and everyday life. The details of this story will concern us in the chapters that follow. For the moment, it is sufficient to note that while the logical use of

<sup>27</sup> The meaning of ‘pure’ in this context is ‘without admixture of anything foreign’ (A11).

<sup>28</sup> On the emergence of Kant’s analytic/synthetic distinction and its relevance to his critique of traditional metaphysics, see Anderson 2015.

reason aims at finding highest principles of *cognition* (principles from which more specific principles and cognitions can be derived, but that cannot themselves be derived from more general ones), the real use of reason, according to Kant, aims at finding first principles of *things* (fundamental sources, elements, or causes that ground other, less fundamental things but are not themselves grounded in anything more fundamental). Kant calls both the highest principles of cognition and the first principles of things “unconditioned” (A307/B364).<sup>29</sup> While the logical use of reason aims at bringing systematic *unity* to the manifold of our cognition by subsuming it under ‘unconditioned principles,’ the real use of reason aims at cognizing the plurality of objects of cognition by tracing them to their unconditioned grounds or conditions (A322/B379). In this way, the aims of the real use of reason strictly parallel those of its logical use. Both are interested in ‘the unconditioned’: the logical use of reason in unconditioned cognition, the real use of reason in unconditioned aspects of reality. (We will investigate the relation between the two uses of reason in more detail in the following chapters.)

### 1.1.6 *Pure Speculative Reason*

In sum, we can see that *pure speculative reason* – the aspect of reason that is central to the Rational Sources Account – is the faculty of gaining a priori cognition (merely purported or genuine) in a purely discursive way (that is, through mere rational thinking based on a priori concepts, principles, and inferences, independently of sensibility and intuition). Its aim is cognizing a domain of objects, and its hoped-for result is knowledge of its unconditioned conditions.<sup>30</sup> To this real use of reason corresponds its logical use, which consists in drawing mediate inferences and which aims at bringing unity to our cognitions, transforming them into a coherent system of knowledge.

In distinguishing between a ‘logical’ and a ‘real’ use of reason, Kant is building on the traditional distinction between rational insight into principles and logical reasoning. However, he is doing so in a way that radically transforms this distinction, since reason can only be discursive for Kant, even in its ‘real’ employment. Kant famously insists that there are two independent, irreducible ‘stems’ or ‘sources’ of human cognition, namely sensibility and

<sup>29</sup> I am assuming here that the term “the unconditioned” at A307/B364 refers to an unconditioned cognition. I defend this reading in Section 2.2.2.

<sup>30</sup> In the first *Critique*, Kant uses the term ‘pure speculative reason’ only in the B-Preface (1787). Kant distinguishes between the “speculative and practical use of pure reason” at A841/B869, but it seems that only after the ‘discovery’ (implicit in the *Groundwork*, but fully explicit only in the second *Critique*) that pure reason can be ‘practical’ (in the specific sense of determining the will through the motive of respect for the Moral Law) did Kant feel the need to distinguish clearly between ‘pure reason’ in general and ‘pure speculative reason’ in particular.

understanding (in the widest sense, including reason), the former of which is intuitive and the latter discursive (A15/B29; A50/B74). Thus, the only intuitive representations we can have are sensible representations. Since reason, for Kant and for the philosophical tradition, is a non-sensible, purely intellectual capacity, this means that rational insight into first principles, if it is possible at all, can only be discursive. Whether human beings have a capacity for a priori insight into principles is clearly relevant to the question of whether metaphysics is possible. Since Kant does not explicitly argue for his fundamental distinction between sensibility and understanding, it can seem that he rules out the possibility of rational intuition by definition. But, as I have argued elsewhere, the *Critique of Pure Reason* contains the materials for an argument for the claim that human intuition can only be sensible and, by implication, that reason must be completely discursive (Willaschek 2015). Since our main aim here is not to discuss Kant's critique of metaphysics but to interpret his attempt to trace metaphysical speculation to its sources in reason, this is not the place to pursue this issue further.

## 1.2 Kant's Conception of Metaphysics

Since the beginnings of Neo-Kantianism, interest in Kant's *Critique of Pure Reason* has mainly concentrated on its more 'constructive' parts, the Transcendental Aesthetics and Analytics, which contain Kant's account of space and time as pure forms of human intuition (and the philosophy of mathematics that is based on it), his defense of a priori knowledge of nature as a necessary condition of the possibility of experience, and his account of experience, according to which the latter is the result of the human mind's activity of synthesizing a sensible manifold in accordance with the categories of the understanding. But while the Neo-Kantians tended to emphasize these aspects of Kant's philosophy at the cost of Kant's metaphysical interests,<sup>31</sup> it has long been recognized that Kant's overarching concern in the first *Critique* is not with science, mathematics, or possible experience, but rather with the possibility of *metaphysics*.<sup>32</sup> It is this issue that, according to the A-Preface, motivates the whole project of a critique of pure reason (Axii); the entire book, according to the B-Preface, is "a treatise on the method" of metaphysics (Bxxii). As Kant famously puts it in a letter to Marcus Herz, the *Critique of Pure Reason* is a "metaphysics of metaphysics" (10:269), that is, a metaphysical theory about

<sup>31</sup> See e.g. Cohen 1871.

<sup>32</sup> This 'metaphysical turn' after Neo-Kantianism is often associated with Heimsoeth 1924, Wundt 1924, and Heidegger 1929.



the possibility of metaphysics.<sup>33</sup> Such a theory is called for, according to Kant, because of the dismal state of the metaphysics of his time. First, there are no successes in metaphysics that can compare to those in mathematics and the sciences (Bxiv); rather, metaphysics presents itself as a “battlefield of endless quarrels” (Aviii). But second, and more importantly, Kant thinks that these quarrels are not due to the failure of individual philosophers; instead, they have their source in reason itself, which, when it ventures beyond the realm of possible experience, entangles itself in fallacies and contradictions. Thus, in order to investigate whether metaphysics can be a respectable scientific enterprise at all, one must first subject pure reason itself to critical scrutiny to determine the conditions, and limits, of its successful use (Axii).

In this section, I will first outline Kant's conception of metaphysics and its sub-fields, based on the Architectonic section of the first *Critique* (Section 1.2.1). I will then briefly indicate which parts of metaphysics, so conceived, can become successful sciences according to Kant, and which cannot (Section 1.2.2).<sup>34</sup>

### 1.2.1 Kant's Conception of Metaphysics in the First Critique

In the section entitled ‘The Architectonic of Pure Reason’ (A832/B860–A851/B879), Kant provides a classification of the ‘rational sciences,’ that is, sciences based on reason alone. ‘Reason’ here means the faculty of a priori cognition, including mathematical cognition (A835/B863). A ‘science’ is a body of knowledge that has the form of a ‘system,’ that is, that has a hierarchical structure, criteria of completeness, and a set of ends, all of which are given a priori in the ‘idea’ of that science (A832–3/B860–1). Kant starts by distinguishing philosophy and mathematics as the two ‘rational sciences (a priori),’ where philosophy is ‘cognition from concepts’ and mathematics ‘cognition from the construction of concepts’ (namely construction of mathematical objects in pure intuition; A837/B865). In this way, Kant makes it clear from the outset that philosophy is a purely discursive enterprise, in contrast to mathematics, which essentially involves intuition.<sup>35</sup>

<sup>33</sup> Thus, the *Critique of Pure Reason* is an essay in what is now called ‘metametaphysics’ and is the topic of a growing body of recent literature (e.g. Chalmers et al. 2009; Tahko 2015).

<sup>34</sup> For an overview of the development of Kant's ‘critical’ account of metaphysics from 1775 to the *Progress* essay, see Ludwig 2017.

<sup>35</sup> One might object that philosophy cannot be purely discursive, according to Kant; after all, the Transcendental Aesthetic discusses space and time as pure forms of intuition, and the Transcendental Analytic essentially refers to the possibility of experience. But this does not show that philosophy itself relies on intuition in the way mathematics does (or in the way the empirical sciences do). Intuition, according to Kant, is not part of philosophical reflection itself, although of course it can be one of its topics (e.g. in the Transcendental Aesthetic).

Within philosophy, Kant then distinguishes between philosophy of nature and of morals, on the one hand, and between “pure” philosophy (or “cognition from pure reason”) and “empirical philosophy” (or “rational cognition from empirical principles”), on the other (A840/B868). Metaphysics is ‘pure philosophy,’ that is, purely rational cognition from concepts, which in turn consists of a ‘critique’ (of pure reason), which is merely preparatory, and metaphysics proper, which Kant characterizes as “the system of pure reason (science)” and as “the whole (true as well as apparent) philosophical cognition from pure reason in systematic interconnection” (A841/B869). Metaphysics, according to Kant, is therefore characterized by two main features: its ‘pureness,’ that is, its *discursivity* and complete independence from experience and even from a priori intuition, and its *systematicity* (which it shares with all other sciences).<sup>36</sup> Both features follow directly from Kant’s claim that metaphysics is ‘cognition from pure reason,’ since pure reason is not just discursive and a priori but also, as we have seen, essentially oriented toward systematic unity.<sup>37</sup>

In this way, Kant insists that metaphysics is not defined by the generality of its principles, as metaphysicians from Aristotle to Wolff and Baumgarten had claimed (metaphysics as “the science of the first principles of human cognition,” A843/B871, as Kant puts it, effectively quoting §1 of Baumgarten’s *Metaphysica*). The generality of metaphysical theses is not sufficient to distinguish them from general empirical claims, and their aprioricity alone cannot distinguish them from mathematics (A843–4/B871–2). Rather, we can have a clear conception of metaphysics only by recourse to its “sources” (A837/B865), “origin” (A844/B872), or “seat” in pure reason: “Thus all pure a priori cognition, by means of the special faculty of cognition in which alone it can

Philosophical knowledge is *reflective* knowledge (Bix; 9:12) based on reason and understanding, not knowledge of objects distinct from these faculties themselves. While the understanding brings discursive unity to the manifold of sensible intuitions, reason brings discursive unity to the manifold of judgments. The a priori knowledge provided by reason and understanding is reflective knowledge about the principles that govern the spontaneous activity of the mind in bringing about these two kinds of unity. I think that this is what Kant means when he says at the end of the Transcendental Dialectic that “pure reason is in fact concerned with nothing but itself” (A680/B708); see also Rohs 1987. (I owe this point to a conversation with Clinton Tolley.)

<sup>36</sup> Emphasizing only the ‘pure’ aspect, Kant characterizes metaphysics in the B-Preface as “a wholly isolated speculative cognition of reason that elevates itself entirely above all instruction from experience, and that through mere concepts (not, like mathematics, through the application of concepts to intuition)” (Bxiv).

<sup>37</sup> Like all sciences, metaphysics requires an ‘idea’ that provides its a priori ‘architectonic’ structure and an end at which it is oriented (A832/B860). In the case of metaphysics (which Kant identifies with philosophy “in a genuine sense”; A850/B878), the end is moral and ultimately consists in “universal happiness” (*allgemeine Glückseligkeit*; A851/B879), which I take to be happiness in accordance with universal moral principles, i.e. the ‘highest good’ (A810/B838). For discussion of the role of the ends in science and philosophy, see Gava 2014 and Sturm (in press).

have its seat, constitutes a special unity, and metaphysics is that philosophy which is to present that cognition in systematic unity" (A845/B873).

Within metaphysics, Kant further distinguishes between "metaphysics of nature" and "metaphysics of morals" (A841/B869). Just as metaphysics is 'cognition from pure reason,' so the distinction between metaphysics of nature and of morals is based on the distinction between the two fundamental uses of pure reason, speculative and practical. In the first *Critique*, Kant sets aside the latter and focuses exclusively on the former. Metaphysics of nature results from the "speculative . . . use of pure reason," contains "pure principles from mere concepts," and is concerned with the "theoretical cognition of all things" (A841/B869). It is this "metaphysics of speculative reason . . . which has customarily been called metaphysics in the narrower sense" (A842/B870) – and which is the kind of metaphysics with which the Rational Sources Account is concerned.

Within speculative metaphysics, Kant further distinguishes, first, between ontology (which is concerned with all *possible* objects) and physiology of nature (which is concerned with all *given* objects, that is, with all objects that actually exist) and, second, between immanent and transcendent metaphysics, where the former is concerned with objects of possible experience (namely either with material or with thinking beings) and the latter with non-empirical objects (such as God or the world as a whole).<sup>38</sup> Kant sums up his division of metaphysics as follows:

Accordingly, the entire system of metaphysics consists of four main parts. 1. Ontology. 2. Rational Physiology. 3. Rational Cosmology. 4. Rational Theology. The second part, namely the doctrine of nature of pure reason, contains two divisions, *physica rationalis* and *psychologia rationalis*. (A846/B874)<sup>39</sup>

This list, however, does not make explicit that unlike rational cosmology and theology, which are transcendent disciplines, and *physica rationalis*, by which Kant means an immanent discipline (developed in Kant's *Metaphysical Foundations of Natural Science* from 1786), *psychologia rationalis* can be either

<sup>38</sup> It may seem surprising that rational theology is subsumed under a metaphysics of *nature*, given that God himself is not part of nature. Kant's reason seems to be that God is here considered as a ground of nature (as a being that is in 'connection with,' but 'above' nature; A846/B874).

<sup>39</sup> This division is closely related but not quite identical to the traditional distinction between *metaphysica generalis* and *metaphysica specialis*, the former of which was traditionally identified with ontology, while the latter consisted of the three specific branches of metaphysics, namely rational psychology, cosmology, and theology – a distinction that structures the text on which Kant's lectures on metaphysics were based: Baumgarten's *Metaphysica* (§2; see Gawlick and Kreimendahl 2011: xlvi). Although Kant never mentions this widely used distinction explicitly in his published writings, only referring to it in his notes and lectures (e.g. *Refl.* 4851, 18:8–9; 28:617), it clearly underlies the structure of the Transcendental Logic of the first *Critique*, which discusses ontology in the Analytic and rational psychology, cosmology, and theology in the Dialectic.

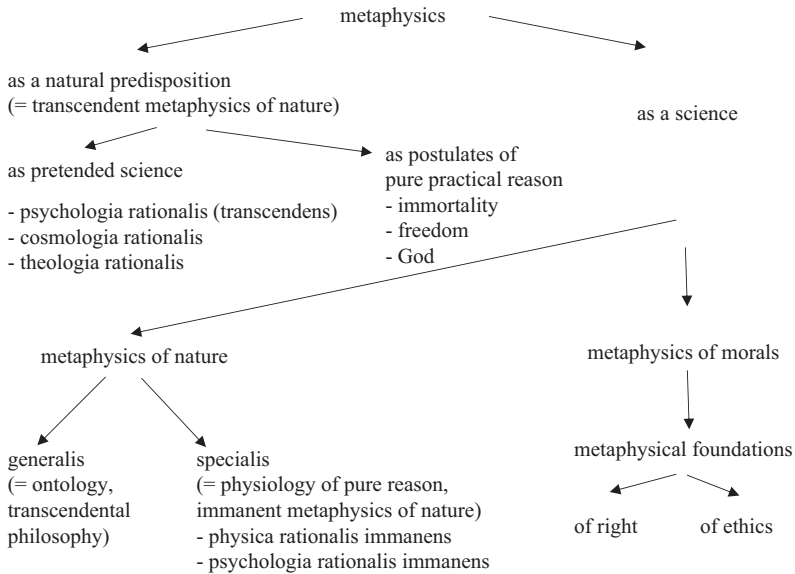


Figure 1.1 The branches of metaphysics according to Kant.

immanent *or* transcendent, since the soul can be considered either as an empirical object or as supersensible.<sup>40</sup> In other words, what is missing in Kant’s list is the kind of *psychologia rationalis transcendens* that is the topic of the Paralogisms chapter. If we add transcendent psychology to the list and combine Kant’s own division at A846/B874 with his distinctions between (a) metaphysics as a natural propensity and as a science and (b) the metaphysics of nature and of morals, and also take into account that the objects of transcendent metaphysics (immortality, freedom, God) ultimately receive some positive epistemic status as ‘postulates of pure practical reason’ (see Postscript), we arrive at the division of Kantian metaphysics shown in Figure 1.1.

When Kant claims in the A-Introduction that reason inevitably raises metaphysical questions that it cannot answer (because they go beyond possible experience; Avii), he is thinking not of questions concerning the foundations of morality (which Kant answers in the *Groundwork*, the second *Critique*, and the *Metaphysics of Morals*), nor of *ontological* questions (which he answers in

<sup>40</sup> This was at least Kant’s view at the time of the A-edition of the first *Critique*, where he seems to have thought that psychology could be a science in the strict sense based on immanent metaphysical (rational) principles. In his *Metaphysical Foundations* (1786), Kant came to deny this (4:471), without, however, revising the relevant passages of the B-edition of the first *Critique*. On Kant’s views on psychology as an empirical science, see Sturm 2001 and Kraus (in press).

the Analytic of Concepts of the first *Critique*), nor of questions concerning *immanent* metaphysics (which he answers in the Analytic of Principles and in the *Metaphysical Foundations*). Instead, he is thinking of questions of *transcendent* metaphysics – that is, transcendent psychology, cosmology, and theology. In fact, Kant tends to identify metaphysics with *transcendent* metaphysics (Bxx; B7; B395n). It is metaphysics exclusively *in the transcendent sense*, going beyond the bounds of possible experience, that is at stake in Kant's Rational Sources Account.

### 1.2.2 *Kant on the Prospects and Failures of Metaphysics*

When Kant defines metaphysics proper as the “system of pure reason” – that is, as systematic, purely a priori, and discursive – he adds that it comprises both “true as well as apparent” cognition (A841/B869). In this section, I want briefly to survey the different parts of metaphysics that emerged in the previous section, with an eye to the question of whether the cognitions they contain are ‘true’ or merely ‘apparent.’ In this way, we will get a better, if still preliminary, understanding of where the Rational Sources Account is located within Kant's overall conception of metaphysics.

Even though metaphysicians before Kant had presented their theories as scientific, Kant claims that they did not succeed, which becomes apparent from the contradictions both within their theories and between different theories that are equally well argued for – contradictions that cannot be resolved by means of the same kind of first-order rational reflection on the subject matter at hand (e.g. objecthood, the soul, the world, God). What has to be asked is the second-order question of how cognition based on this kind of reflection is possible in the first place, which is the project of a critique of pure reason (see Kant's letter to Marcus Herz, February 12, 1772; 10:129–35). Metaphysics without critique (that is, merely on the basis of our ‘natural predisposition’ to ask metaphysical questions) necessarily ends in fallacies and contradictions. Therefore, metaphysics as (successful) science is possible *only* on the basis of a critique of pure reason.

Now this critique of pure reason has two fundamental results concerning the possibility of metaphysics as a science: first, the claim that cognition requires both intuition and concepts, and second, the distinction between appearances and things in themselves.

First, *any* human cognition (in the demanding sense of ‘cognition’ relevant here) requires contributions from both sensibility and understanding, intuitions and concepts (A51/B75).<sup>41</sup> Put crudely, while only sensible intuitions relate

<sup>41</sup> Kant also uses ‘cognition’ in a wider sense for any representation that purports to represent something beyond a merely subjective state; see Watkins and Willaschek 2017 on Kant's different uses of the term ‘cognition.’

directly to (possible) objects, only the concepts of the understanding allow us to cognize these objects, that is, to attribute properties to them. As Kant famously puts it, “[w]ithout sensibility no object would be given to us, and without understanding none would be thought. Thoughts without content are empty, intuitions without concepts are blind . . . Only from their unification can cognition arise” (A51/B75). But from this it seems to follow that metaphysical cognition is impossible, given that it would have to be purely discursive (‘cognition from mere concepts’).

That this cannot be Kant’s considered view is obvious from the fact that he takes himself to have shown that metaphysics *is* possible once we distinguish properly between appearances and things in themselves (e.g. Bxvii–xviii). This is the *second* fundamental result of the *Critique of Pure Reason* concerning the possibility of metaphysics. How to understand the distinction between things in themselves and appearances is a matter of some debate. For now, it will suffice to say that appearances are (all and only) empirical objects, of which Kant claims to have shown that all their empirical properties depend on the way human beings represent them (specifically, on the way we represent them in our human forms of sensible intuition, space and time). By contrast, things in themselves are not represented in space and time and thus do not depend on our forms of intuition. Since, according to Kant, things in themselves cannot be given in human sensibility (because they are not represented in space and time), we cannot have cognitions of them – we can only have cognition of appearances.

Combining the two claims – cognition requires sensible intuition, the objects of which, for human beings, are mere appearances – Kant argues that metaphysical cognition *is* possible if it concerns the “conditions of the possibility of experience” (A158/B197). Because objects of possible experience are mere appearances, they depend on the forms in which we represent them (space, time, categories). For instance, Kant argues that the principle ‘Every alteration has a cause’ is a condition of the possibility of experience, which means that we can know a priori that it holds for all objects of (humanly) possible experience (B232–4).

These are weighty philosophical claims, and this is not the place to discuss their credentials. But if we accept them for the moment, it follows that we can know a priori that empirical objects are spatiotemporal, that they must exhibit the structure of objecthood specified by the categories, and that they must conform to fundamental principles that follow from applying the categories to objects in space and time. Besides causation and substantiality, these principles concern the quantitative and qualitative structure of empirical reality, as well as the modal properties of our judgments about it. For our purposes, it will not be necessary to go into the details of these a priori principles and Kant’s arguments for them. What matters is only that all of these claims and principles are metaphysical in Kant’s sense.

More specifically, the categories outline the structure of all possible objects of human experience and are thus the basis for a critical *ontology* (which, according to Kant, includes the system of claims that follows analytically from the categories and the a priori concepts that can be defined on their basis, the so-called predicables; A79–82/B105–8).<sup>42</sup> Applying this structure to objects of possible experience leads to what Kant, in the Architectonic, calls 'immanent metaphysics.' Kant's immanent metaphysics includes the 'principles of the pure understanding' (A158/B197–A235/B287) and his account of matter and physical forces in the *Metaphysical Foundations*.

However, from Kant's limitation of human cognition to the objects of possible experience it also follows that any attempt to develop *transcendent metaphysics* into a science is doomed to fail. After all, transcendent metaphysics would have to consist in a priori claims about non-sensible objects – objects that cannot be given in human intuition, such as an immaterial soul, the world as a whole, and God. It is one of Kant's aims in the Transcendental Dialectic to show in detail how and why the supposed proofs of transcendent metaphysics fail (the other aim, as we have seen, being to argue for the Rational Sources Account, that is, to reconstruct the ways in which rational thinking naturally leads to transcendent metaphysics). But even before the Transcendental Dialectic begins, it is clear according to Kant that transcendent metaphysics can never be a successful science. As Kant writes, looking back at the Transcendental Dialectic: "The outcome of all dialectical attempts of pure reason . . . confirms *what we have already proved in the Transcendental Analytic*, namely that all the inferences that would carry us out beyond the field of possible experience are deceptive and groundless" (A642/B670; emphasis added). Thus, for Kant, the failure of transcendent metaphysics already follows from its very definition as a kind of metaphysics that transcends the boundaries of possible experience.<sup>43</sup>

Of the major branches of metaphysics distinguished by Kant, this leaves only the metaphysics of morals, or practical metaphysics, to be discussed. Even though Kant does not apply the distinction between immanent and transcendent metaphysics to the practical realm, it seems appropriate to think of the theory Kant develops in the *Groundwork*, the *Metaphysics of Morals*, and parts of the second *Critique* as being analogous to the immanent

<sup>42</sup> Kant sometimes suggests that ontology (or its 'proud name') is to be *replaced* by the results of the Transcendental Analytic (A247/B303), but then, as we have seen, he nevertheless includes it among the sub-disciplines of a scientific metaphysics (A846/B874).

<sup>43</sup> In Chapter 9, I will argue that this result does not presuppose Kant's transcendental idealism, with its distinction between appearances and things in themselves, but only his claim that there can be no cognition without sensible intuition, which, although no doubt contentious, can be defended quite independently of transcendental idealism.

metaphysics of nature unfolded in the Principles chapter of the first *Critique* and the *Metaphysical Foundations*. In both cases, Kant is concerned with metaphysical claims about objects of possible experience – in the moral case, claims about human beings and their moral obligations. Even though the epistemic status of these claims is a matter of some debate (since Kant never develops his ‘practical’ epistemology in much detail), it is clear that Kant himself thought of this project as (a) metaphysical (abstracting from all empirical knowledge about human beings, based only on reason; 4:388–9) and (b) successful (e.g. 6:216–17). The metaphysical foundations of morality, further subdivided into those of right and ethics, constitute a scientific theory in Kant’s sense (which is why Kant begins the Preface of the *Groundwork* with a reflection on the division of the sciences, one of which turns out to be the metaphysics of morals; 4:387–8). By contrast, Kant’s so-called doctrine of the postulates concerns ‘transcendent’ objects, namely God, immortal souls, and freedom of the will, which can never become objects of science. Nevertheless, Kant insists that something analogous to a science is possible even here. Although we cannot have theoretical cognition or knowledge of God, the soul, or freedom, the claims that God exists, that our souls are immortal, and that our wills are transcendentally free can be rationally warranted – although not in the mode of knowledge (*Wissen*), but only in the mode of belief or rational faith (*Glaube*). (We will briefly return to Kant’s practical metaphysics and how its success is related to the failure of speculative metaphysics of nature in the Postscript.)

### 1.3 Conclusion

Kant’s conceptions of reason and metaphysics are made for each other. While speculative reason aims at systematic knowledge of the unconditioned, transcendent metaphysics is rational and systematic cognition of the soul, the world, and God (which, as we will see later, are unconditioned objects). This may seem to trivialize the Rational Sources Account, according to which the use of reason necessarily leads into metaphysical speculation. One may suspect that the Rational Sources Account follows directly from Kant’s extremely demanding, speculation-prone conception of pure reason. But in fact, this suspicion is unfounded. It is true that, given Kant’s conception of pure speculative reason, the Rational Sources Account follows more or less trivially. But Kant’s claim is that metaphysical questions raised by pure reason can be traced back to features inherent in ‘universal human reason.’ Thus, the argumentative work for the Rational Sources Account consists not in showing that the use of pure speculative reason, unsurprisingly, leads to metaphysical speculation, but rather in showing that ordinary employments of reason have a tendency to take us to a speculative use of reason. As we will see in Chapters 4



and 5, Kant's argumentative strategy is to start from the logical use of reason – the use of inferences in non-metaphysical investigations, both in everyday life and in the sciences – and to show that it naturally takes us to the real use of reason, which in turn leads to metaphysical speculation. But first, we must look more closely at Kant's account of the logical (Chapter 2) and real uses (Chapter 3) of reason.

## 2 The Logical Use of Reason and the Logical Maxim

---

In a central but opaque passage from the Introduction to the Transcendental Dialectic, Kant writes:

[T]he proper principle of reason in general (in its logical use) is to find the unconditioned for conditioned cognitions of the understanding, with which its unity will be completed. But this logical maxim cannot become a principle of *pure reason* unless we assume that when the conditioned is given, then so is the whole series of conditions subordinated one to the other, which is itself unconditioned, also given (i.e., contained in the object and its connection). (A307–8/B364)

Kant here distinguishes between (at least) two principles: the ‘Logical Maxim’ that requires us to “find the unconditioned for conditioned cognitions of the understanding” and a principle that says “when the conditioned is given, then so is the whole series of conditions subordinated one to the other, which is itself unconditioned, also given.” A few lines later, Kant refers to the latter principle as the “*supreme* principle of pure reason” (A308/B365; emphasis added); we will call it the ‘Supreme Principle.’ While the Logical Maxim is concerned with conditioned and unconditioned *cognitions*<sup>1</sup> (‘the unconditioned for conditioned cognitions’), the Supreme Principle is about conditioned and unconditioned *objects* (‘contained in the object and its connection’). In other words, **while the Logical Maxim belongs to reason in its *logical* use, the Supreme Principle is part of its *real* use (Section 1.1.5)**. According to Kant, the Logical Maxim is a principle of reason that guides rational thinking in general and scientific inquiry in particular; considered on its own, however, it does not have any metaphysical implications. As Kant puts it, it is merely a “subjective law of economy for the provision of our understanding” (A306/B362). By contrast, the Supreme Principle is a metaphysical principle about objects and relations between them. Given that there are conditioned objects (for instance, caused objects

<sup>1</sup> Cognitions, in the relevant sense, are representations that determine given objects by attributing general properties to them (Watkins and Willaschek 2017). Cognitions in this sense can, but need not, have propositional form. Those cognitions on which the logical use of reason works, however, are propositional (namely premises and conclusions of inferences).

and wholes conditioned by their parts), the Supreme Principle implies that something ‘unconditioned’ exists.

Now Kant’s claim in the cited passage – a version of the Rational Sources Account (see Introduction, Sections 0.2 and 0.3) – is that the Logical Maxim cannot ‘become’ a principle of pure reason ‘unless we assume’ the Supreme Principle. While this formulation is somewhat cryptic – we will be able to explain it in detail only in Chapter 4 – the context makes it clear that **Kant’s central point is this: if we accept the Logical Maxim as a guiding principle of our cognitive economy, we must also accept the Supreme Principle (and, so it seems, thereby accept the existence of something unconditioned).**

This raises a number of questions that we will try to answer in this chapter and the three chapters that follow. What are conditioned cognitions, what are unconditioned cognitions, and why is it an a priori principle of reason (‘in its logical use’) to search for the latter once the former are given? What does it mean to say that objects are conditioned or unconditioned, and why is a complete series of conditions itself unconditioned? And finally, how can Kant claim that we must accept the Supreme Principle (assuming we must follow the Logical Maxim), even if the Supreme Principle takes us beyond the realm of experience into metaphysical speculation (and ultimately into antinomies and delusions)? Answering these exegetical questions will shed further light on Kant’s rich and highly complex account of human reason, a first sketch of which was given in the previous chapter. In the present chapter, we will begin by looking more closely at Kant’s account of the logical use of reason and its guiding principle, the Logical Maxim. Even though some aspects of this account may seem outdated, its central elements will turn out to be philosophically important and plausible, even from a contemporary perspective.

## 2.1 The Logical Use of Reason

As a first approximation, the Logical Maxim can be formulated as follows: ‘If there is a cognition that is conditioned, seek the cognition that is its condition. If this condition is itself conditioned, seek *its* condition, and so on, until you find some unconditioned cognition.’<sup>2</sup> According to Kant, the Logical Maxim is the “proper principle of reason in general (in its logical use)” (A307/B364). So what is the logical use of reason? As we have already seen,

<sup>2</sup> On the reading of ‘the unconditioned’ in the Logical Maxim as ‘unconditioned condition,’ see Section 2.2.2. Kant’s own formulation requires us to *find* the unconditioned for every conditioned cognition (A307/B364), but I take this to mean that we should *seek* the unconditioned. **‘Finding,’ strictly speaking, would imply that the unconditioned exists, which is something the Logical Maxim cannot and need not presuppose.** Kant more appropriately speaks of ‘seeking’ (*gesucht werden*) instead of ‘finding’ earlier in the same sentence that introduces the Logical Maxim (A307/B364).

Kant distinguishes the logical from the real use of reason in the Introduction to the Transcendental Dialectic:

As in the case of the understanding, there is in the case of reason a merely formal, i.e., logical use, where reason abstracts from all content of cognition, but there is also a real use . . . The first faculty has obviously long since been defined by the logicians as that of drawing inferences mediately. (A299/B355)

So reason in its logical use is characterized by two features: (1) it is formal in that it abstracts from all content, and (2) it draws mediate inferences (i.e. inferences that require two or more premises).

### 2.1.1 Formality

The logical use of reason is formal, according to Kant, in that it “abstracts from all content of cognition” (A299/B355). Here, Kant is implicitly referring back to the Introduction to the Transcendental Logic, where “logic” is defined as the “science of the rules of understanding in general” (A52/B76) and “general logic” (as opposed to subject-sensitive logics) is said to “abstract from all content of the cognition of the understanding and of the difference of its objects, and has to do with nothing but the mere form of thinking” (A54/B78). A little later, Kant glosses abstraction from the “content of cognition” as abstraction from “any relation of it to the object”; general logic “considers only the logical form in the relation of cognitions to one another” (A55/B79). Transcendental logic, by contrast, does not abstract from all content of cognition (A55/B80) in that it considers “the origin, the domain, and the objective validity” (A57/B81) of pure cognitions a priori, and thus their relation to objects. This precisely parallels the way in which Kant distinguishes between the logical and the real use of reason insofar as the former, like *general* logic, is said to abstract from all content and not to concern objects, while the latter, like *transcendental* logic, is concerned with “the origin of certain concepts and principles” (A299/B355).

While there has been debate about Kant’s distinction between general and transcendental logic, as well as the sense in which the former is formal while the latter is not (e.g. Wolff 1995: 197–231; MacFarlane 2002; Tolley 2012), it seems safe to say that the logical use of reason is formal in that it concerns only the (logical) relations among our cognitions and not their relation to objects. Since the logical use of reason is said to consist in the drawing of inferences, the logical form in question is that of syllogisms and other rational inferences. For instance, the inference ‘All humans are mortal; Caius is human; therefore, Caius is mortal’ has the same logical form as the inference ‘All bodies are alterable; the earth is a body; therefore, the earth is alterable,’ and this logical form can be expressed by using concept variables: ‘All *A* are *B*; *C* is *A*;

therefore, *C* is *B*.' That the logical use of reason abstracts from all content of cognition means that the validity of the inference does not depend on which concepts we fill in for the variables and how their objects (e.g. humans and bodies, Caius and the earth) differ from one another. All that matters are the logical relations among the cognitions in question.<sup>3</sup>

### 2.1.2 Rational Inference

Kant explains inference in general as the truth-preserving progression from one sentence, or set of sentences, to another (A303/B360) and contrasts the mediate inferences of reason with the immediate inferences of the understanding. By the latter he means inferences that do not require a minor premise because the conclusion follows from the major premise alone. According to traditional syllogistic logic, for instance, both 'Some humans are mortal' and 'Some mortals are human' follow directly from 'All humans are mortal' (A303–4/B360; 9:118–19).

Kant's most general characterization of rational inferences (*Vernunftschlüsse*)<sup>4</sup> is in terms of 'universal rules,' 'conditions,' and 'subsumption': "An inference of reason is the cognition of the necessity of a proposition through the subsumption of its condition under a given universal rule" (9:120). A 'rule,' or universal judgment, according to Kant, consists of a condition and an assertion (9:121). For instance, in 'All humans are mortal,' being human is the condition under which being mortal is universally asserted. We can derive the claim 'All scholars are mortal' by subsuming its condition ('scholar') under the condition of the major premise 'All humans are mortal' ('human'), which requires the minor premise 'All scholars are human.' Thus, the universal principle of all rational inferences is: "*What stands under the condition of a rule also stands under the rule itself*" (9:120). In this sense, the conclusion is *conditioned* by the premises.<sup>5</sup>

<sup>3</sup> The fact that Kant characterizes the logical use of reason in terms of formality and abstraction from content shows that he thinks of it as part of general logic rather than transcendental logic. By contrast, it is the *real* use of reason that is characterized as 'transcendental,' that is, as concerning the 'origin' of our concepts and principles and their 'relation to objects.' This explains why Kant tends to use both 'real' and 'transcendental' as exchangeable antonyms of 'logical' in the Introduction to the Transcendental Dialectic; e.g. A299/B355–6; A305–6/B362–3). See also Chapter 1, note 23.

<sup>4</sup> See Chapter 1, note 4, for the choice of 'rational inference' (or, equivalently, 'inference of reason') as a translation of Kant's term '*Vernunftschluss*.'

<sup>5</sup> Within rational inferences, Kant distinguishes between categorical, hypothetical, and disjunctive inferences according to the logical form of their major premises (A304/B361; 9:122) – a distinction that will turn out to be important at subsequent levels of Kant's Rational Sources Account. While categorical inferences have categorical sentences as their major premises (e.g. 'All men are mortal'), the major of a hypothetical inference is hypothetical (e.g. 'If there is an alteration, then there is a cause of the alteration') and that of a disjunctive inference is

Inferences of reason differ from those of the understanding not just in requiring an additional premise but also in requiring a *specific cognitive activity*, namely that of ‘determining’ the subject term of the minor (‘scholar’) by applying to it the predicate term of the major (‘mortal’) (A304/B360–1). This activity differs from concept formation and concept application in individual judgments (which are tasks Kant attributes to the understanding), from subsuming the particular under the universal (faculty of judgment), and from immediate inferences (understanding). The logical use of *reason* in rational inference is a specific cognitive achievement analogous to, but different from, the ‘synthesis’ involved in concept application. In applying a concept to an empirical object, we must ‘synthesize’ a manifold of sense impressions into the cognition of the empirical object according to the marks contained in that concept (A105; B137). In drawing an inference from two premises, we must ‘synthesize’ the assertion of the major premise (e.g. ‘being mortal’) with the condition of the minor (e.g. ‘being a scholar’) into the conditioned cognition expressed by the conclusion (e.g. ‘All scholars are mortal’). And just as concept application is discursive in that it involves a succession of (at least logically) distinct steps (‘going through, taking up, and combining’ the given manifold sensations in accordance with the multiple marks contained in the concept; A77/B102), so rational inference is discursive, since it involves the successive synthesis of elements (the premises) that are given prior to their being synthesized.<sup>6</sup>

Kant’s theory of syllogistic inference is complex and goes back at least to his 1762 essay on the Aristotelian syllogistic figures (*False Subtlety*). There is no need to engage it here in greater detail (see e.g. [Stuhlmann-Laeisz 1976](#); [Malzkorn 1995](#); [Dahlstrom 2015a](#)). There is only one more aspect that will become important in what follows, namely that syllogisms can be combined into *polysyllogisms*, that is, into series of syllogisms where the conclusion of one inference is a premise of another (9:133–4). For instance, we can use the conclusion in our example (‘All scholars are mortal’) as the major premise from which to derive ‘Some philosophers are mortal’ with the aid of the minor premise ‘Some philosophers are scholars.’ Or we can derive the major premise in our original example (‘All humans are mortal’) as a conclusion from ‘All living beings are mortal’ and ‘Humans are living beings.’<sup>7</sup> Kant calls a syllogism (and more generally a rational inference) whose conclusion serves as a premise of a further syllogism a ‘prosyllogism’; he calls the further

disjunctive (e.g. ‘The world is either finite or infinite,’ where the disjunction is supposed to be an exclusive dichotomy).

<sup>6</sup> See Section 0.2 for the concept of ‘discursivity.’

<sup>7</sup> But cf. [Klimmek 2005: 26](#), who argues – contra Kant – that in the categorical inferences that interest Kant, the prosyllogism concerns not the major but the minor.

syllogism, which takes the conclusion as a premise, an ‘episyllogism’ (9:134).<sup>8</sup> While rational inferences as such exhibit the discursivity of reason, polysyllogisms also give expression to its iterative character.

### 2.1.3 Inferential Concatenation

As Kant points out, rational inferences can be used for two different purposes: (1) to deduce propositions that are not yet elements of our body of cognition or knowledge and (2) to deduce propositions that we already know to be true from more general propositions that are known to be true (A304/B361).<sup>9</sup> The inferences and the relevant cognitive activity are the same in both cases, but they are employed for different tasks. The former task was ridiculed by modern philosophers (such as Bacon) in their critiques of scholastic philosophers, who supposedly restricted their endeavors to deducing conclusions from premises according to the rules of Aristotelian logic (Kneale and Kneale 1984: ch. 5). This made it impossible to attain new insights, the critics argued, because syllogisms can only make explicit what is implicitly contained in the premises.<sup>10</sup> This criticism leaves the second task of syllogistic reasoning unaffected, however. This is the task of articulating the inferential structure of a given body of cognition or knowledge – of explicating what can be derived from what.

Like Wolff and Meier (see Section 1.1.4), Kant sees the main purpose of reason and rational inferences as lying not in gaining new knowledge but in expressing the logical relations between various parts of our body of cognition. The picture Kant suggests in the Introduction to the Transcendental Dialectic looks like this: we start with the ‘manifold of cognition’ supplied by the understanding, which consists of a priori principles (such as ‘Every alteration has a cause’), empirical laws (such as the laws of physics), and empirical generalizations (such as ‘Dolphins are mammals’).<sup>11</sup> Next, we combine as

<sup>8</sup> In the case of a hypothetical inference, we can either derive its minor premise as a conclusion from another hypothetical inference or use its conclusion as a minor premise in a further inference (18:222); analogously for disjunctive inferences.

<sup>9</sup> A third use consists in deriving conclusions (either already known to be true or not) from hypothetically assumed premises; see Section 2.1.4. As Kant makes clear in the Appendix to the Transcendental Dialectic, we can inductively justify a general principle by deriving consequences from it that are already known to be true (A647/B675; see Section 4.2.2).

<sup>10</sup> This objection to the traditional syllogistic logic laid out in Aristotle’s *Organon* motivated Bacon’s project of a ‘New Organon,’ which was supposed to provide heuristic principles for the study of nature (*Novum Organon*). Kant admired Bacon (Bxii; 9:32) and took the motto of the B-edition of the *Critique of Pure Reason* (Bii) from the very preface in which Bacon rejects traditional logic (“dialectica”) because it “perpetuates error” instead of “opening the way to truth” (Bacon, *Instauratio magna*, Praefatio).

<sup>11</sup> It is unclear to me whether Kant would also include singular empirical judgments such as ‘Flipper is a dolphin.’ Since he is ultimately interested in a scientific system of cognitions, it seems more plausible that they are not included.

many of these cognitions as possible into syllogisms by looking, for each cognition, for more general cognitions from which it can be derived. Finally, we combine inferences into *polysyllogisms* with the aim of finding, for each cognition, the *most general* cognition from which it can be derived through a series of syllogisms.

Consider an example Kant uses in a similar context: the proposition that all bodies are alterable (A330/B387).<sup>12</sup> Like all general judgments, it consists of a ‘condition’ (is a body) and an ‘assertion’ (is alterable). We now “seek whether the assertion of this conclusion is not to be found in the understanding under certain conditions according to a universal rule” (A304/B361). So what we look for is a universal cognition with the same assertion (alterable) but with a different condition, such as ‘Everything composite is alterable.’ Kant continues: “Now if I find such a condition and if the object of the conclusion can be subsumed under the given condition then this conclusion is derived from the rule *that is also valid for other objects of cognition*” (A304–5/B361). We subsume ‘being a body’ under the condition of the rule that everything composite is alterable and thus arrive at the minor premise ‘All bodies are composite,’ which allows us to deduce our original judgment as a conclusion. In this way, we have subsumed a more particular cognition under a more general one (‘also valid for other objects of cognition’) and have thus taken a step toward unifying our body of cognitions. Moreover, we have given a specific kind of explanation of our original judgment by having shown that bodies are alterable *because* they are composites. We can call this kind of explanation ‘inferential explanation,’ which consists in the recognition that some seemingly isolated fact turns out to be an instance of something more general (which is also valid for other cases). Inferential explanation is the specific task of reason, since reason is the faculty of cognizing “the particular in the general through concepts” (A300/B357).

But this need not be the end of this process, since we can now ask whether there are even more general cognitions from which to derive our premises. Let us assume that there are: we can derive, say, ‘Everything composite is alterable’ from ‘Everything composite has parts’ and ‘Everything that has parts is alterable.’ And perhaps we can derive ‘All bodies are composite’ from ‘Everything extended is composite’ and ‘All bodies are extended.’ In this way, we have derived our original cognition, ‘All bodies are alterable,’ from a number of more general cognitions. That they are more general means that they are

<sup>12</sup> At A330–1/B387, Kant uses the example differently in that he does not assume that we already know that the conclusion is true, so that we can come to know its truth only by deriving it from more general premises. By contrast, at A304/B360–1 Kant assumes that in searching for premises the conclusion is already known ‘through the understanding,’ as the final sentence of the section makes clear. This paragraph and the next are adapted from Willaschek 2008.



‘also valid for other objects of cognition’ besides bodies; adding appropriate minor premises, we can also derive ‘All gases are alterable,’ ‘All souls are unalterable,’ etc.

If we repeat this process of inferential concatenation of our cognitions until all cognitions are included, we ideally arrive at a hierarchically ordered system of cognitions, with the most general ones at the top and the most particular ones at the bottom.<sup>13</sup> Thus, Kant can ascribe to reason, in its logical use, the task of bringing “the great manifold of cognition of the understanding to the smallest number of principles (universal conditions), and thereby to effect the highest unity of that manifold” (A305/B361). The unity in question is what Kant calls the “unity of reason” (*Vernunftseinheit*), which he distinguishes from the synthetic unity of the understanding (A302/B358–9). While the latter results from synthesizing a sensible manifold in accordance with the categories, the former results from bringing the manifold cognitions of the understanding under ‘principles.’

#### 2.1.4 *System and Science*

The Introduction to the Transcendental Dialectic makes it sound as if the unity of reason consisted in nothing more than a hierarchical ordering of cognitions according to their generality, which is achieved by placing them in a net of rational inferences. But when Kant returns to the idea of the unity of reason after more than 300 pages in the Appendix to the Transcendental Dialectic, a richer picture emerges. There, the unity of reason is said to presuppose

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)

Thus, the proper task of reason is to achieve “the *systematic* of cognition, i.e. its interconnection based on one principle” (ibid.) – that is, turning our cognitions into a *system* (which Kant defines as “the unity of the manifold cognitions under one idea,” A832/B860).<sup>14</sup>

<sup>13</sup> ‘Ideally’ because there are many obstacles, including the fact that there might not be a unique way of ordering our cognitions into a hierarchical system.

<sup>14</sup> Kant’s conception of systematicity and its relevance to his conception of science, on the one hand, and to his own philosophical ‘system,’ on the other, is a complex topic which I cannot adequately address here. On the systematic character of Kant’s own philosophy, see the contributions in Fulda and Stolzenberg 2001, as well as Guyer 2000 and Henrich 2001. On systematicity and science, see e.g. Kitcher 1994; Neiman 1994: ch. 2; Sturm 2009: 129–82. On the specific way in

Systems are characterized as having a guiding *idea* (such as the idea of the soul, in the case of psychology; A671–2/B699–700), as being *complete* (which is guaranteed by the guiding idea), as having *unity* (under principles or ‘necessary laws’), and as having an *a priori structure* (A832/B860). In the Appendix, Kant attributes the search for systematic unity to the logical use of reason:

[S]ystematic unity or the unity of reason of the manifold of the understanding’s cognition is a *logical* principle, in order, where the understanding alone does not attain to rules, to help it through ideas, simultaneously creating unanimity among its various rules under one principle (the systematic), and thereby interconnection, as far as this can be done. (A648/B676)

By identifying ‘systematic unity’ with the ‘unity of reason’ (which according to the Introduction is the constitutive goal of the logical use of reason; A305/B361; A307/B364), Kant makes explicit that the latter involves more than ordering given cognitions according to their generality, namely *an idea* that determines the way in which the parts of the system are supposed to hang together. **Kant thinks of the logical use of reason, in accordance with his teleological conception of cognitive faculties, as being goal directed.** It consists not in drawing inferences for their own sake, as it were, but in doing so with the aim of achieving systematic unity among a given body of cognition, where the logical place of each cognition within the system is defined a priori by a guiding idea.

By emphasizing systematicity, Kant aligns the logical use of reason with striving for scientific knowledge, since science, according to Kant, is characterized by its systematic structure (A832/B860). But scientific knowledge must be not only systematic but also certain (4:468). Accordingly, in the Appendix Kant distinguishes between the apodictic and the hypothetical use of reason (apparently as two varieties of the logical use of reason), where the former proceeds from universal premises that are “certain,” while the latter assumes premises that are not certain but “problematic” in order to see whether consequences that are themselves certain follow from them (A646–7/B674–5).<sup>15</sup>

which systematicity as a necessary condition of science applies to philosophy, see Gava 2014. The differences between Kant’s own system of philosophy as it emerges in the Doctrine of Elements (including the Transcendental Dialectic) and the system sketched in the Architectonic are discussed in Goy 2007. For an overview of the different meanings of ‘system’ and their functions in Kant’s work, see Dahlstrom 2015b.

<sup>15</sup> The relation between the concepts of science and certainty in Kant is more complex than my remarks might suggest, since Kant distinguishes not only different kinds of certainty (e.g. 4:468: “apodictic” and “empirical”; 9:70: “rational” and “empirical”) but also different kinds of sciences (e.g. 9:72: “rational” and “historical”) and a narrower and wider conception of science (e.g. 4:468). For instance, Kant seems to regard only “empirical certainty,” not “apodictic” certainty, as a requirement for empirical or ‘historical’ sciences, which therefore

By invoking the concepts of certainty and inductive confirmation, the distinction between the apodictic and the hypothetical use of reason adds an important *epistemological* dimension to the picture of the logical use of reason painted in the Introduction to the Transcendental Dialectic. While there it looked as if the logical use of reason was only concerned with the logical articulation of the body of our cognitions, it now emerges that it is also concerned with the epistemic status of individual cognitions and relations of epistemic justification between different cognitions. This aspect of the logical use of reason can come into view only once it is made explicit that its ultimate aim is systematicity, and thus science.

Against this picture of the logical use of reason, it might be objected that, strictly speaking, the logical use only concerns the logical articulation of a body of cognitions or knowledge (as portrayed in the Introduction to the Transcendental Dialectic), while epistemological concerns come in only once we also consider the objects of our cognitions and thus make real use of reason (as portrayed in the Appendix).<sup>16</sup> However, as the citations above show, in the Appendix Kant himself attributes the concern with systematicity and certainty not to the real or transcendental but to the logical use of reason (A646–8/B674–6), which also fits the fact that certainty, for Kant and his contemporaries, is a *logical* feature of cognitions (9:65–6; Meier, *Vernunftlehre*, §§9, 29).

Nevertheless, there is a valid point in the objection, which concerns the fact that **the logical and the real use cannot always be as neatly separated as the Introduction to the Transcendental Dialectic may suggest**. This can be brought out by distinguishing between the logical use of reason *in abstracto*, that is, in abstraction from all content, and *in concreto*, that is, applied to a specific content. The latter characterization may appear to be contradictory, since the logical use of reason is *defined* by its abstraction from content (A299/B355; see Section 2.1.1). But in fact, there is no such contradiction. Compare the logical schema of a syllogism of the *Barbara* figure: ‘All *A* are *B*; all *B* are *C*; therefore, all *A* are *C*’ with the inference ‘All bodies are composites; all composites are alterable; therefore, all bodies are alterable,’ as that inference features in a particular person’s body of cognitions. While the logical schema expresses the logical use of reason *in abstracto* (as is done for instance in a logic textbook), the latter clearly also involves the logical use of reason, but this time *in concreto*, as applied to a specific subject matter. Despite its application to a concrete content, this is a case of the logical use of reason in that the validity of the inference does not depend on its content, but only on its logical form. Perhaps we can say that Kant’s account in the Introduction of the Transcendental Dialectic is (primarily)

do not count as sciences in the strictest sense (4:468). On Kant’s conception of science and the role of certainty, see Sturm 2009: 146–53.

<sup>16</sup> Thanks to Eric Watkins for pressing me on this.

meant to capture the logical use of reason *in abstracto*, while the account in the Appendix considers the logical use of reason *in concreto* and therefore emphasizes its epistemological aspects more strongly.<sup>17</sup>

## 2.2 The Content of the Logical Maxim

Given his account of the logical use of reason, Kant arrives at the Logical Maxim in two steps (A306–7/B363–4). First, he reminds the reader that rational inferences do not refer to objects (or to intuitions that refer to objects) directly; instead, they refer to objects indirectly, through the understanding. This means that the ‘material’ on which the logical use of reason is exercised are (general) cognitions of the understanding. Second, in a syllogism we subsume a condition under a general rule:

Now since this rule is once again exposed to this same attempt of reason, and the condition of its condition thereby has to be sought (by means of a prosyllogism) as far as is possible, we see very well that the proper principle of reason in general (in its logical use) is *to find the unconditioned for conditioned cognitions of the understanding*, with which its unity will be completed. (A307/B364; emphasis added)

Kant goes on to call this principle a “logical maxim.”<sup>18</sup> Before we can ask whether Kant is correct to attribute this principle to the logical use of reason – that is, whether it is plausible to assume that the Logical Maxim should guide the way in which rational beings organize their body of cognitions or beliefs – we must be clear about its content. Kant’s formulation (‘find the unconditioned for conditioned cognitions of the understanding’) raises a number of questions and allows for several different readings.

### 2.2.1 Conditioned Cognition

First, it is unclear what Kant means by ‘conditioned cognitions of the understanding.’ In particular, what does it mean to say that a cognition is

<sup>17</sup> We can also consider the real use of reason in *abstracto* (that is, in abstraction from the logical use of reason), as a list of concepts and principles that arise from pure reason alone, and in *concreto*, as an inferentially structured system of (real or merely purported) metaphysical knowledge. This shows that it would be a mistake to think of the logical and real uses of reason as necessarily distinct. When we reason about a specific subject matter, we can distinguish between form and matter and consider the form in abstraction from the matter (e.g. in doing formal logic), and we can also consider the matter independently of its logical form (e.g. in writing a metaphysics textbook without the proofs). But of course, both typically *occur* together.

<sup>18</sup> Kant distinguishes maxims, which are “subjective principles,” from laws, which are “objective” (e.g. 4:400 n.; 5:19). While in his ethics Kant is interested in practical maxims that are principles of action, the Logical Maxim is a theoretical principle in that it concerns (theoretical) cognition. As Kant explains in the Appendix to the Transcendental Dialectic, “maxims of reason” are “subjective principles that are taken not from the constitution of the object but from the interest of reason in regard to a certain possible perfection of the cognition of this object” (A666/B694).

*conditioned*? Given that Kant introduces the Logical Maxim by appealing to the idea that syllogisms subsume a condition under a general rule and that we must find ‘the condition of the condition’ by means of prosyllogisms, **the most obvious answer is that a cognition is conditioned if it can be derived from more general cognitions by means of a syllogism.** Thus, ‘All bodies are alterable’ is conditioned in this sense, since it can be inferred from ‘Everything composite is alterable’ and ‘All bodies are composites.’ Let us call this ‘*inferentially conditioned cognition.*’

inferential  
conditioning

A stronger reading emerges if we assume that **a conditioned cognition not only can but *must* be inferred from other cognitions *in order to be cognized.*** The passage in which Kant uses the inference to ‘All bodies are alterable’ as his example suggests such a reading:

Thus suppose I arrive at the proposition ‘All bodies are alterable’ *only* by beginning with the more remote cognition . . . ‘Everything composite is alterable,’ and go from this to . . . ‘Bodies are composite’; and then from this finally to . . . : ‘Consequently, bodies are alterable’; then I arrive at a cognition (a conclusion) through a series of conditions (premises).<sup>19</sup> (A330/B387; emphasis added)

If we take this as our model, then a cognition is conditioned if it can be cognized *only* by being inferred from a set of premises, which serve as its conditions.<sup>20</sup> Such a reading faces a difficulty, however, since it seems to imply that all statements that can be cognized in other ways than by being inferred from general premises (for instance, all empirical cognitions supplied by the understanding) count as *unconditioned*, which surely is not what Kant means.

This difficulty can be circumvented, however, if we assume that the epistemic status at stake is stronger than that of ‘ordinary’ cognition provided by the understanding. **For instance, a cognition might count as conditioned (in the relevant sense) if the only way for it to constitute *scientific knowledge* is by being derived from general premises.** As we have seen, **Kant mentions an “apodictic use” of reason, which consists in inferring conclusions from premises that are epistemically certain (A646/B674).** In this way, certainty can be transmitted from premises to conclusions that, independently of the inference, are not already certain. It seems plausible to read this idea back into Kant’s account of the logical use of reason in the Introduction and to say that ‘conditioned cognition’ includes cognitions whose certainty (or, more generally, positive epistemic status) depends on their being derivable by logical

<sup>19</sup> Note that here, as in various other places, Kant calls the premises themselves, and not the subject terms of general sentences, ‘conditions.’

<sup>20</sup> See also A331/B388: “we cannot reach it [the cognition] by means of reason except at least on the presupposition that all members of the series on the side of conditions are given . . . because only under this presupposition is the judgment before us possible *a priori.*”

inference from other cognitions (as their conditions). We can call these cognitions *epistemically conditioned*.<sup>21</sup> The conditioning relation is epistemic in that it concerns the epistemic status required for a cognition to be part of a system of scientific cognitions, such as knowledge or certainty.<sup>21</sup>

Both conditioning relations between cognitions, inferential and epistemic, involve kinds of explanation (in the widest sense of the term): while an inferentially conditioned cognition is partly explained by being cognized as a particular instance of a more general principle (what I earlier called ‘inferential explanation’; Section 2.1.3), the epistemic status of an epistemically conditioned cognition is explained by deriving it from the epistemic status of some other cognition.<sup>22</sup> Another way of making this point is by saying that conditioned cognitions raise ‘why’ questions of a specific kind (‘Why is it the case that  $p$ ?’, ‘Why is it certain that  $p$ ?’) and that the logical use of reason consists in providing answers to these questions by deriving the cognitions in question from other, more general ones (‘ $p$  is the case because  $q$ , of which  $p$  is a specific instance’; ‘ $p$  is certain because it can be derived from  $q$ , which is certain’).<sup>23</sup> As we will soon see, even though inferential and epistemic conditions do not necessarily coincide, Kant is particularly interested in the case of cognition that is *neither* inferentially *nor* epistemically conditioned (that is, both inferentially and epistemically *unconditioned*). This will become clear once we ask what Kant, in the context of the Logical Maxim, can mean by ‘the unconditioned.’

### 2.2.2 Unconditioned Cognition

In the A-version of the *Critique of Pure Reason*, the term ‘the unconditioned,’ which is ubiquitous throughout the Transcendental Dialectic, occurs for the

<sup>21</sup> A third possible reading is suggested by what Kant says two paragraphs after introducing the Logical Maxim, where he mentions “objects of a possible experience, whose cognition and synthesis are always conditioned” (A308/B365). Kant does not elaborate on this, but here ‘conditioned cognition’ might mean a cognition that is ‘transcendentally’ conditioned by something being given to us in sensibility and by the conditions of sensibility (space and time) and of the understanding (categories). However, this reading does not fit Kant’s characterization of the logical use of reason. After all, the logical use of reason connects cognitions (which are representations, not objects) with other cognitions by means of inferences. But the transcendental conditions of cognition are not premises from which the cognitions in question can be derived. They are not logical but *real* conditions, concerning dependence relations between one type of entity (representations that qualify as cognitions) and another type of entity (the mind with its a priori forms). While I do not want to rule out the possibility that Kant may have wanted the Logical Maxim to apply to transcendentally conditioned cognitions, I will not pursue this possibility any further.

<sup>22</sup> The connection between reason and explanation is emphasized in, e.g. Rohlf 2010. For a contemporary account of reason that links reason to explanation, see Schafer 2017.

<sup>23</sup> The close connection between talk of conditions and ‘why’ questions is emphasized by Proops 2010, who refers to Baumgarten’s *Metaphysica* (§14) and Kant’s logic lectures (24:921) but does not distinguish between ‘logical’ and ‘real’ conditions in this context.

first time in the formulation of the Logical Maxim (“to find the unconditioned for conditioned cognitions of the understanding”, A307/B364).<sup>24</sup> Since Kant does not explain what he means by this term, he seems to assume that the reader can gather its meaning from the context. Things are further complicated by the fact that in the German original, the term ‘*das Unbedingte*’ (with a capital ‘U’) is a noun, which makes it difficult (but not impossible) to read it as an elliptical expression for ‘unconditioned cognition.’ This might suggest that Kant is talking about an object called ‘the unconditioned’ (or, somewhat less bewilderingly, an unconditioned object). But such a reading is difficult to square with Kant’s view that the logical use of reason is concerned with *cognitions* (in abstraction from their content) and not, like the real use of reason, with objects. Moreover, Kant contrasts this ‘unconditioned’ with ‘conditioned *cognition*,’ a term which suggests the possibility of ‘unconditioned cognition.’ In fact, that Kant uses ‘unconditioned’ as a noun here does not exclude the possibility that the unconditioned at stake is ‘the unconditioned among our cognitions,’ that is, some unconditioned cognition. Since this seems to be the reading that makes best philosophical sense, I will assume that the Logical Maxim directs us to find some *unconditioned cognition* for each piece of conditioned cognition.<sup>25</sup>

But what does it mean to say that some cognition is *unconditioned*? This of course depends on what one means by ‘conditioned cognition.’ If we mean *inferentially* conditioned cognition, the unconditioned is a principle (universal premise) from which other cognitions can be derived but that cannot in turn be derived from other premises. If we mean *epistemically* conditioned cognition, the unconditioned is a principle that has some positive epistemic status (e.g. certainty) that is not derived from anything else. We can call the former unconditioned an *inferentially* first principle and the latter an *epistemically* first principle. I now want to suggest that only principles that are both *inferentially and epistemically* first principles are unconditioned in the sense required by Kant’s conception of an (ideal) system of scientific knowledge.<sup>26</sup>

inferentially vs  
epistemically first  
principles

<sup>24</sup> In B, the term “the unconditioned” is also used in the Preface (Bxx).

<sup>25</sup> Against this reading, it might be objected that Kant in fact never speaks of ‘unconditioned cognitions.’ This is correct, but it does not rule out the reading suggested here. First, as mentioned in the text, Kant does speak of ‘conditioned cognitions,’ which seems to require, as a logical contrast, the possibility of cognitions that are not conditioned. Thus, at least in this indirect sense, Kant does speak of ‘unconditioned cognitions.’ And second, in the *Jäsche Logic*, Kant defines ‘principles’ as judgments (cognitions) that are (1) self-evident (‘immediately certain’) and (2) not derivable from more general premises (not ‘subordinated’ to others) (9:110). Thus, they are neither epistemically nor inferentially conditioned (in the sense explained in the previous subsection), which means that Kant acknowledges that there is a specific type of unconditioned cognition, even if he does not use that expression for them.

<sup>26</sup> Anticipating a discussion in Chapter 3, a comparison between the ‘logical’ and the ‘real’ use of reason might suggest that first principles are not the only candidates for unconditioned cognitions because, on the side of the real use of reason, there are *two* ways in which something can

Note that inferentially first principles need not necessarily be epistemically first principles; conversely, cognitions whose positive epistemic status does not depend on other cognitions may not be the most general ones. These two things will come apart, for instance, in empiricist epistemologies such as Locke's, since the epistemically first principles will be something like sensations or perceptions, which obviously are not inferentially first principles (that is, most general cognitions). In some rationalist epistemologies, by contrast, inferential and epistemic priority will tend to coincide: for Baumgarten, for instance, the principle of non-contradiction is both an inferentially first principle, since it cannot be derived from any other cognition, and an epistemically first principle, since it is self-evident, which means that its epistemic status (certainty) is underived (*Metaphysica*, §7). Might it be the case that Kant did not explicitly distinguish between inferential and epistemic conditions of cognition because he accepted such a rationalist conception of science?

Not quite. First, even though Kant does not seem to distinguish between inferentially and epistemically conditioned cognitions in the Introduction to the Transcendental Dialectic, he does draw what is essentially the same distinction in a different context. Thus, in the introductory passages to the section entitled "System of all Principles of Pure Understanding," Kant points out:

A priori principles bear this name not merely because they contain in themselves the grounds of other judgments, but also because they are *not themselves grounded in higher and more general cognitions*. Yet this property *does not elevate them beyond all proof*. For although this could not be carried further objectively, . . . yet this does not prevent a proof from . . . subjective sources . . . from being possible, indeed even necessary, since otherwise the proposition would raise the *greatest suspicion of being a merely surreptitious assertion*. (A148–9/B188; emphasis added)

So here Kant distinguishes between an a priori principle's not being grounded in 'higher and more general cognitions,' that is, its status as inferentially unconditioned, and its being in need of a proof, that is, its status as being epistemically conditioned. While Kant's further distinction between objective and subjective proofs complicates matters in ways we need not discuss here, it is clear that he allows for the possibility that inferential and epistemic unconditionality can come apart.

be unconditioned: either by being a first (unconditioned) condition or by being the (possibly infinite) totality of (conditioned) conditions (Section 3.3.4). Similarly, a cognition might be unconditioned either by being an inferentially and/or epistemically first principle or by being the totality of inferential/epistemic conditions. In the latter case, the system of cognitions as a whole would count as unconditioned, even if none of its constituent conditions would. Even though Kant does not seem to consider this possibility, it fits the holistic aspect of Kant's account of systematicity.



Second, we saw before that the logical use of reason, according to Kant, aims at a scientific system of cognitions, which is characterized by unity, completeness, a priori structure, and certainty. But Kant distinguishes between empirical and rational certainty (9:70–1). While the former is based on experience (either one's own or that of others), the latter is a priori and accompanied by the "consciousness of necessity." Empirical certainty based on testimony is called "historical certainty" (9:71). Accordingly, Kant distinguishes between "historical" sciences and "sciences of reason" (9:72). But this distinction between empirical and rational certainty is not exclusive: "our cognitions can concern objects of experience and the certainty concerning them can still be both empirical and rational at the same time, namely, insofar as we cognize an empirically certain proposition from principles *a priori*" (9:71). For instance, we may have empirical reasons to believe that all bodies are alterable, but we can also derive this proposition as a conclusion from a priori premises (e.g. 'All bodies are composite' and 'Everything composite is alterable').

This suggests the following picture. In empirical/historical sciences, inferentially first principles and epistemically first principles can (and presumably will) come apart. But in rational sciences such as metaphysics, even empirical cognitions (that is, cognitions that *can* be arrived at empirically), if they are admitted at all, must be 'cognized from' epistemically first principles by being derived from them, because only in this way can they be rationally certain. In this case, the epistemically first principles will also be inferentially first principles. The 'consciousness of necessity' that accompanies rational certainty derives from the fact that the cognitions in question are either 'principles a priori' or (directly or indirectly) inferred from them. (Recall that an inference, according to Kant, is the "cognition of the necessity of a proposition," 9:120.)

So the reason why Kant did not explicitly distinguish between inferentially and epistemically (un)conditioned cognition in the Introduction to the Transcendental Dialectic may have been that, in the context of his discussion of the logical use of reason, he was only interested in 'rational sciences,' in which certainty (or, more generally, positive epistemic status) is transmitted 'downward,' from principles that are *both inferentially and epistemically* first principles to other cognitions that are *both inferentially and epistemically* conditioned. Empirical cognitions may find a place in such a rational system, but they will count as rationally certain only insofar as they can be inferred from a priori principles.

If this is correct, it brings Kant's conception of the logical use of reason very close to Meier's account of reason. As we saw in Section 1.4, Meier, on whose book Kant based his logic lectures, posits the following:

In a demonstration from reason all grounds the proof is based on must be completely certain (§§193, 204); hence they are either demonstrable or indemonstrable

(§313). In the first case, they in turn must be proven. Consequently, a proof will not become a demonstration [from reason] until I arrive at indemonstrable grounds only. (*Auszug*, §318)

In other words: reason requires certainty (which can be either rational or empirical; §157), and cognitions are certain if they are either demonstrable (that is, epistemically conditioned and proven) or indemonstrable (that is, epistemically unconditioned).<sup>27</sup> The chain of proofs of demonstrable cognitions must ultimately end in indemonstrable ones. Meier also claims that syllogistic reasoning serves to transform a manifold of cognitions into a systematic unity (e.g. §413), which he calls ‘science’ (§434). A body of cognition (*Lehrgebäude*, doctrinal edifice) is a ‘system’ (*systema*; §104); if it is presented according to the ‘synthetic method,’ according to Meier, all cognitions can be derived from one supreme principle, so that the edifice is characterized by unity, coherence, and completeness (§431).<sup>28</sup>

Thus, Kant found in Meier’s logic textbook a model for his own account of a scientific system of cognitions and of the logical use of reason, just as he found in the logic textbooks of his time (A70–1/B96) a model (albeit an imperfect one) for his account of judgment and the “logical use of the understanding” (A67/B92). Kant’s overall strategy in the *Transcendental Logic* – in both the *Analytic* and the *Dialectic* – is to take the uncontentious ‘logical use’ of the understanding (forms of judgment) and reason (rational inference) as a ‘guiding thread’ or ‘clue’ (*Leitfaden*; A66/B91) for finding their corresponding ‘real’ or ‘transcendental’ use (the categories in the case of the understanding, the Supreme Principle and the transcendental ideas in the case of reason). For this purpose, it is important for Kant that his account of the logical use of reason is not just his own invention but is based on the standard logic of his time.<sup>29</sup> At the same time, the parallels between Meier’s account of rational proof, syllogistic inference, and science and Kant’s account of the logical use of reason confirm that Kant indeed seems to have thought of the logical use of reason as aiming at a scientific system of cognition, and thus as including not just inferential but also epistemological conditioning relations.

<sup>27</sup> Note that indemonstrable or self-evident cognitions do not have to be thought of as *self-justifying* (in which case they would not be epistemically *unconditioned*, but *self-conditioned*); rather, they can be regarded as not standing in need of justification. We will see in Section 3.4 that Kant thinks of at least some ‘real’ conditioning relations as irreflexive.

<sup>28</sup> For a comparison between Kant’s conception of science and systematicity and those of his predecessors, in particular Wolff and Meier, see Hinske 1991; Baum 2001; Sturm 2009: 139–46; and Gava (in press), who seem to agree that, despite many similarities, only Kant requires that a system be based on a guiding ‘idea.’

<sup>29</sup> This is why Kant explicitly mentions “the logicians” when he introduces the concept of a logical use of reason (A299/B355), just as he does after presenting the table of judgments (A70–1/B96).

In sum, we can see that the distinction between conditioned and unconditioned cognition, as it features in the Logical Maxim, has to be understood against the background of Kant's conception of rational science, which is a hierarchically structured body of cognitions, all of which are linked by inferential relations (expressed in syllogisms and other rational inferences). Ideally, in such a system there is only one supreme principle from which all others can be deduced. This principle is *both* inferentially and epistemically unconditioned; that is, it is the most general principle under which all other cognitions can be subsumed and at the same time self-evident, so that it can transmit rational certainty to all other cognitions (by means of rational inference). When Kant attributes to the logical use of reason the task of bringing "the great manifold of cognition of the understanding to the smallest number of principles" (A305/B361), the aim is such a system of scientific knowledge. Therefore, 'conditioned cognition' should be understood to consist of cognitions that are *either* inferentially *or* epistemically conditioned, and unconditioned cognition as consisting of principles that are *both* most general (inferentially unconditioned) *and* self-evident (epistemically unconditioned).

### 2.2.3 *The Logical Maxim: The Full Formulation*

Our reflections so far result in the following version of the Logical Maxim:

Logical Maxim      If there is some piece of cognition that is inferentially or epistemically conditioned, seek the cognitions that are its inferential or epistemic conditions, respectively. If these conditions are themselves inferentially or epistemically conditioned, seek the cognitions that are *their* inferential or epistemic conditions, and so on, until you find cognitions that are both inferentially and epistemically unconditioned.<sup>30</sup>

Kant does not tell us to whom the Logical Maxim is addressed – to each individual rational being or to all rational (human) beings collectively. Correspondingly, it is unclear whose 'manifold cognitions' are to be systematized by following this maxim – those of an individual thinker or those of rational (human) beings in general. As Kant repeatedly emphasizes in the *Critique of Pure Reason*, science in general, and philosophy in particular, is a communal enterprise; it can be successful only if many individuals and even many generations of researchers cooperate (e.g. A820/B848; A856/B884). On the other hand, this cooperation must ultimately consist in the efforts of individual

<sup>30</sup> For an alternative reading, cf. Klimmek 2005: 23, who suggests three specific versions of the Logical Maxim, one for each of the three relational categories (categorical, hypothetical, disjunctive); see also Grier 2001: 119–21.

people. Therefore, I will assume that the Logical Maxim addresses each rational being individually but concerns not their private and accidental sets of beliefs but rather those that can possibly be integrated into a rational system of scientific knowledge that can be shared by all rational beings.<sup>31</sup> Moreover, the aim at which the Logical Maxim is directed is not one that any individual is supposed to realize alone. Rather, the task of the individual in following the Logical Maxim is to *contribute* to realizing (or approximating) a complete system of scientific knowledge.<sup>32</sup> Put differently, nothing in what Kant says suggests that the Logical Maxim requires that each individual person should try to transform their own body of cognitions into a scientific system.

Two further restrictions should be noted. First, the Logical Maxim is not a categorical imperative: it does not direct us to find or look for unconditioned cognitions come what may, like moral imperatives (which on Kant's view hold without exception). Rather, Kant points out repeatedly that the speculative interest of human reason is only 'conditional' (5:5; 5:142; 8:139; 9:87), which means that we are rationally required to pursue it only when doing so is morally permitted and pragmatically feasible.<sup>33</sup> Second, the Logical Maxim directs us to turn our cognitions into a unified system, but only "as far as possible" (A307/B364; "as far as this can be done," A648/B676). So the aim is to *approximate* such a system, not necessarily to realize it fully (A647/B675). The general idea behind that maxim is to turn a body of cognitions (alternatively: representations, beliefs, doxastic attitudes, statements), through a series of steps (inferences, justifications), gradually into an inferentially structured whole, so that ideally each element of that whole receives some positive epistemic status (e.g. certainty, justification, knowledge) from occupying a specific place within that whole.

### 2.3 The Logical Maxim, Science, and Universal Human Reason

According to Kant, the Logical Maxim is "the proper principle of reason in general (in its logical use)" (A307/B364). Like other "maxims of reason," it is

<sup>31</sup> These beliefs need not be 'scientific' in our current sense of the term. Rather, they can include all general statements about reality ('cognition of the understanding').

<sup>32</sup> This is meant to be analogous to the way in which, according to Kant, every individual moral agent ought to "advance" (*befördern*) the realization of the highest good (5:114).

<sup>33</sup> It is therefore misleading for Allison to call the Logical Maxim an "intellectual categorical imperative" (Allison 2004: 312; 331). Proops, on the other hand, claims that the Logical Maxim (his prescription 'P') is subjective in depending on a contingent desire: "Should one not wish to proceed rationally in inquiry, one will stand under no obligation" in this respect (Proops 2010: 456). This seems too weak, since according to Kant we ought to proceed rationally in inquiry, whatever our desires. (This follows from Kant's account of imperatives according to which one *ought* to do what is *rational* for one to do; 4:412–13.) Thus, the resulting obligation is conditional on a 'desire' (or 'need'), but the desire is one that is internal to reason and is in this sense necessary.

a “subjective” principle in that it is taken “from the interest of reason in regard to a certain possible perfection” of our cognition (A666/B694). To this ‘subjective’ Logical Maxim there corresponds an ‘objective’ principle – the Supreme Principle – that applies not to cognitions but to objects and that not only requires us to *strive* for the unconditioned but also positively asserts its existence (under the assumption that something conditioned exists). As we will see, Kant questions the objective validity of the Supreme Principle, but he does not question the ‘subjective’ validity of the Logical Maxim, if by that we mean that the Logical Maxim normatively guides the way rational beings (qua scientists) organize their body of cognitions (or beliefs). Kant clearly thinks that the Logical Maxim is a *legitimate* principle of reason precisely because it only concerns cognitions (and not their objects) and only requires us to approximate systematic unity (and does not claim that we can fully realize it).

However, the Logical Maxim is supposed to be a “principle of reason in general,” albeit only in its “logical use” (A307/B364). In other words, it is supposed to be valid not only for ‘speculative reason’ but for ‘universal human reason.’ This point is central to Kant’s overall argumentative strategy in defending the Rational Sources Account, according to which metaphysical questions arise from ordinary uses of reason in everyday life (Section 0.2). The transition from the Logical Maxim to the Supreme Principle constitutes the most general of the four levels on which Kant argues for this claim (Section 0.3). This means that Kant’s argument can be fully successful only if the Logical Maxim is a principle not only of speculative reason but also of metaphysically innocent uses of reason that can be attributed to ‘universal human reason’ or ‘reason in general.’ This point, however, may seem to conflict with the fact that the Logical Maxim directs us toward the goal of a scientific system of cognition, which clearly goes beyond what can plausibly be required of rational beings as such. Moreover, Kant’s conception of science, with its emphasis on systematicity and certainty, can seem outdated in a time where most scientists and philosophers of science would deny that scientific theories are hierarchically structured in the way Kant assumes and that scientific knowledge is certain. We must therefore ask whether it is plausible to consider the Logical Maxim as a guiding principle of human reason in general. I think that even though some aspects of Kant’s account of reason and science may no longer be plausible, the claim that some version of the Logical Maxim is valid for rational beings as such can be defended even from a current perspective.

### 2.3.1 *The Logical Maxim and Science*

Let us first turn to the inferential aspect of the Logical Maxim. One problem is that the inferential relations captured by syllogisms and other rational

inferences (that is, relations of conceptual containment and hypothetical reasoning) are not always the most relevant for scientific understanding and explanation. In biology, we may be interested in a taxonomy with the highest genera at the top and the lowest known species at the bottom. But even within biology, theories such as genetics and neurobiology will not easily lend themselves to being represented in this way. Perhaps it is *possible* to represent our current genetic knowledge as a hierarchical system. But not much would be gained by such a representation because it would not capture the explanatory relations between the statements of the theory.

However, this is not to deny that all scientific theories contain both general principles or laws and more specific claims that fall under them, with the latter standing in some form of logical subordination to the former. And perhaps that is enough to validate the Logical Maxim's requirement to search, for any given cognitions, for higher principles. After all, the Logical Maxim does not presuppose, or even claim, that this search will be successful in each case; rather, it requires us to *look* for unconditioned principles that unify our body of cognition.

That the ideal of a hierarchically ordered system of scientific knowledge is still very much alive today can be seen from the fact that many physicists subscribe to the search for a theory that unites quantum mechanics and general relativity theory. Given that such a theory would explain phenomena from the smallest scale (e.g. quarks) to the largest (e.g. galaxies), and presumably all phenomena in between, scientists like Stephen Hawking have claimed that it would be a 'theory of everything' (Hawking 2005).<sup>34</sup> The quest for such a theory is obviously driven by the idea that different pieces of scientific knowledge must ultimately be subsumed under very few general laws from which every aspect of nature can be explained. Perhaps such a system does not have to be structured by inferential relations in the Kantian sense. But it clearly would be hierarchically structured in the sense that the more specific principles and claims are subsumed under more general ones, from which they can be derived (in a sufficiently wide sense of the term). In other words, it would be a theory that brings "the great manifold of cognition of the understanding to the smallest number of principles" (A305/B361), which is just what the Logical Maxim is supposed to achieve. Now such a 'theory of everything' may just be a mirage, given that the sciences are actually highly fragmented and that all attempts to 'reduce,' e.g. biology to chemistry and chemistry to physics seem to face serious objections.<sup>35</sup> But the intuitive appeal of the ideal of a comprehensive theory of nature in which all specific theories are integrated is not undermined by the current fragmented state of scientific research or by the

<sup>34</sup> See Stevenson 2011a for critical discussion from a Kantian point of view.

<sup>35</sup> See e.g. the classic papers Fodor 1974 and Dupré 1983.

admission that this fragmentation is likely to continue. We gain a deeper understanding of particular facts, laws, and theories by relating them to other facts, laws, and theories; in particular, we deepen our understanding by viewing seemingly distinct phenomena as instances of the same underlying principles. (The paradigm for this is Newton's theory of gravitation, which showed for the first time that the same laws that govern falling objects on Earth also govern the movement of the planets.) Kant's Logical Maxim does not claim that the search for this kind of unification and systematicity will always be successful, but only that we should pursue it 'as far as possible.' The current fragmentation of the sciences does not imply that this is not a meaningful goal.<sup>36</sup>

Concerning the epistemic aspect of the Logical Maxim, it must be admitted that very few scientists or philosophers of science working today would consider certainty to be the epistemic standard that scientific theorems must live up to (at least outside mathematics). One reason for this is a lesson from the history of science, which shows that many scientific theorems that once seemed certain later turned out to be false. The Euclidian character of physical space is but one example among many: while Kant thought that it was an a priori truth that space is Euclidean, relativity theory (and many experiments that confirm it) has shown not only that physical space is not Euclidean but, ipso facto, that this claim has never been certain. Given the history of scientific theories, which has borne witness to the overthrowing of many supposedly certain beliefs, and taking into account the empirical character of scientific theories, which implies that any such theory can be falsified by future experience, certainty just does not seem to be the appropriate epistemic standard for science.<sup>37</sup>

Moreover, even if we apply a less demanding epistemic standard (perhaps something like 'empirically better confirmed than all rival theories'), the idea that the epistemic justification of scientific theorems flows from unjustified justifiers in a succession of steps to all other parts of the theory is incompatible with the holistic character of scientific theories, which makes it impossible to determine the epistemic status of individual claims independently of the empirical adequacy of the theory as a whole. As Quine famously puts it: "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body" (Quine 1953: 41). While such epistemological holism may be contentious when applied to the beliefs of

<sup>36</sup> See Philip Kitcher's Kantian account of scientific explanation as maximizing the number of phenomena explained by one explanatory pattern (Kitcher 1994).

<sup>37</sup> This is not to rule out that scientific claims can be a priori in a suitably relativized sense that detaches apriority from (apodictic) certainty and infallibility. For relativized accounts of the a priori, see e.g. Friedman 2001. For a fallibilist reading of Kant's philosophical methodology, see Gava 2016.

individual people, it seems undeniable with respect to current scientific theories, with their complex apparatus of theoretical terms and principles.

But note that even though the downward transmission of epistemic justification through inferential chains ('polysyllogisms') introduces a foundationalist aspect into Kant's account of science and the logical use of reason, due to Kant's emphasis on systematicity it also has a holistic aspect. After all, Kant's point is that the epistemic status of a given cognition (e.g. its status as scientific knowledge) depends on its place in a unified system. The Logical Maxim requires us to look, for *every* epistemically conditioned cognition, for a set of cognitions from which it can be derived (or, more generally, by appeal to which it can be justified), which is just what a holistic understanding of epistemic justification requires. Moreover, Kant himself seems to allow for 'upward' justification in science with his account of a 'hypothetical use of reason,' mentioned earlier, where hypothetically assumed principles are inductively justified by the fact that more specific cognitions one already possesses can be derived from them (A646–7/B674–5; see also Section 4.2.2). In this way, the search for more general principles from which more specific ones can be derived can be part of a holistic conception of epistemic justification in the sciences.<sup>38</sup>

In sum, the idea that there is a rational requirement to look for general principles from which specific cognitions can be derived can be detached from an exclusively foundationalist conception of epistemic justification and the idea that scientific knowledge must be certain. This means that even though Kant's account of science may contain elements that are no longer plausible, this does not undermine the status of the Logical Maxim as the expression of a valid rational requirement.

### 2.3.2 *The Logical Maxim and Universal Human Reason*

But even if one admits that the Logical Maxim expresses a rational requirement, one may object to the claim that it is part of 'reason in general' in the sense that *every* rational human being is required to search for the conditions of her inferentially and epistemically conditioned cognitions. After all, most people are not scientists and thus simply not in the business of transforming their body of cognitions into a unified system. It may be plausible to assume a universal rational requirement to avoid contradictions in one's body of beliefs and cognitive commitments. But it would be absurd to claim that ordinary people stand under an obligation to look for conditions for each and every one of their inferentially and epistemically conditioned cognitions. For instance, if

<sup>38</sup> For a reconstruction of Kant's account of science that emphasizes its holistic aspects, see Gava 2014.



someone who is not a meteorologist believes that summer in Europe is typically warmer than winter, there does not seem to be any rational pressure to look for a more general principle from which to derive, and thereby epistemically justify, the belief in question. Epistemic justification, at least in everyday contexts, mostly follows a default-and-challenge pattern, which means that we count as justified in our beliefs unless we are confronted with relevant challenges to them (see Williams 2001; Willaschek 2007, 2012; Matthiessen 2014). And even if we are challenged, we do not have to climb the ladder of epistemic conditions up to the unconditioned (to something self-evident or otherwise indubitable); we need only appeal to beliefs that are unchallenged in the present context. As Peirce once remarked, it makes no sense “to argue a point after all the world is fully convinced of it” (Peirce 1992: 115). Thus, the Logical Maxim with its requirement to search for conditions for *every* inferentially and epistemically conditioned cognition (even if ‘all the world is fully convinced of it’) does not seem to hold for ordinary people in ordinary situations. If it holds at all, it only applies to scientists (in the widest sense, including philosophers and other people interested in what Kant calls ‘speculation’).<sup>39</sup>

But even if this is correct, it does not undermine the idea that the Logical Maxim is rooted in universal human reason. After all, the Logical Maxim gives expression to three features of the logical use of reason which seem to hold universally: first, its *discursive* or *inferential* character, leading by a series of steps from one claim to another; second, its *iterative* character, allowing the same discursive operation (e.g. syllogistic inference) to be applied to the result of previous instances of that operation; and third, the requirement that this process must come to an end somewhere (*completeness*) (Section 0.2). Each of these features is intuitively appealing and has legitimate applications.

First, rational inference is a powerful tool of thought that can transmit credence and epistemic justification from premises to conclusions. And even where epistemic justification does not take the form of a syllogism, it is often conveyed in a stepwise manner from one belief to another. Second, iteration is also an important cognitive tool. The conclusion of one syllogism can be made the premise of another, leading us further and further in articulating the logical consequences of our beliefs and in transmitting epistemic justification. And again, even where epistemic justification is transmitted not by inference but by other means, we often do ask not merely for justification but also for the justification of the justifying belief. (We ask, for instance, ‘How do you *know* that?’, the answer to which we follow with ‘And how do you know *that*?’). And, third, it is certainly correct that any such series of questions must end

<sup>39</sup> For a helpful discussion of Kant’s own philosophical claims as having an ‘in between’ status between common sense and science, see Ameriks 2001.

somewhere (completion), even if this need not involve principles that are underivable and self-evident, but perhaps only ones that cannot be reasonably unchallenged in the present context. Assuming that discursivity, iteration, and completion are features of universal reason, we can understand the Logical Maxim as grounded in universal reason, since it gives normative expression to these three features.

This still leaves us with the question of why the Logical Maxim, even though it is grounded in universal human reason, nevertheless typically applies not to ordinary people but only to scientists (in the widest sense of the term). An answer can start from the restricted validity of the Logical Maxim, which, as we have seen, is not a categorical imperative but only applies under the condition that no other, more urgent concerns prevent us from following it (Section 2.2.3). For people who are not scientists, this condition may typically not be satisfied. If someone believes that summer is warmer than winter, the reason why they are not required by the Logical Maxim to search for other cognitions from which to derive and justify the belief in question might just be that for them, there will always be other, practical concerns that are more pressing than the rational interest in scientific explanation. Seen from this angle, the Logical Maxim does hold for everyone, but vacuously so for most, since a necessary condition of its making substantive requirements on us and our cognitive activity is not satisfied in most cases.

## 2.4 Conclusion

In this chapter, we have seen how Kant thinks of the Logical Maxim as guiding the ‘logical use of reason.’ The latter consists in the drawing of rational inferences with the aim of unifying our body of cognitions into a system of scientific knowledge. Ideally, in such a system all our inferentially and epistemically conditioned cognitions could be derived by chains of rational inferences (polysyllogisms) from a small number of principles that are both inferentially and epistemically unconditioned. The Logical Maxim directs us to approximate such an ideal system by seeking conditions for our inferentially and epistemically conditioned cognitions. While such a requirement may have some pull only for scientists (in the widest sense of the term), but not for most other people most of the time, the Logical Maxim can nevertheless be plausibly regarded as part of ‘universal human reason’ since it gives expression to three features (discursivity, iteration, and completeness) that characterize rational thinking as such.