1
Transcendental Illusion

1.1 Introduction

According to Kant, human beings, by virtue of their very rationality, are subject to a peculiar kind of systematic illusion, which he terms 'transcendental illusion'. Illusion of this kind is not erroneous belief or false judgment but rather a non-perceptual *seeming* that produces in rational human subjects an inclination to make false judgments and to form false beliefs. More specifically, it produces in these subjects a temptation to embrace a wide range of dogmatic metaphysical positions. The project of 'transcendental dialectic' is the *critique* of this 'dialectical' illusion (A 63/B 88). It aims to alert us to the existence of transcendental illusion, to explain its workings, and, most ambitiously of all, to persuade us that this species of illusion lies behind all serious and sincere theoretically grounded dogmatic metaphysics—the kind of metaphysics that makes claims to know a priori and on theoretical grounds (and with apodictic certainty) a range of claims about the supersensible.

In calling his critique of dialectical illusion transcendental 'dialectic' Kant is intentionally bestowing a new meaning an old word. In the hands of the ancients, he tells us, dialectic was 'a technique of illusion or a sophistical art' (R 5063, 18: 76–7). It was the art—in Kant's view a wholly disreputable one—of *creating* illusion.[1] Kant's own 'transcendental' dialectic, by contrast, is the art of *exposing* transcendental illusion by means of critique. One does so with the intention not of deceiving others but rather of preventing transcendental illusion from deceiving oneself (A 297/B 354).

[1] *Vienna Logic*, 24: 794; *Jäsche Logic*, 9: 17. Kant nods to this older usage when he figures speculative reason as a tricky dialectical illusionist (A 606/B 634). One suspects, therefore, that a further reason why he calls the part of the *Critique* that lays out the arguments of dogmatic speculative metaphysics 'The Transcendental Dialectic' is that he is likening these arguments to the work of such an illusionist. Thus, while transcendental illusion is natural, Kant is prepared to grant that certain of the arguments it encourages might involve contrivance. At the opening of the Transcendental Dialectic Kant characterizes dialectic in yet a third way, namely, as the 'logic of illusion' [*Logik des Scheins*] (A 293/B 349). Here he is nodding—with some subtle wordplay ('*Schein*' versus '*Wahrscheinlichkeit*')—to an alternative tradition according to which dialectic is the logic of probability (*logica probabilium*). Kant, however, alludes to this tradition only to distance himself from it. For details see Giorgio Tonelli, 'Die historische Ursprung der Kantischen Termini "Analytik" und "Dialektik"', *Archiv für Begriffsgeschichte*, 7 (1962), 120–39 at 135–6.

The Fiery Test of Critique: A Reading of Kant's Dialectic. Ian Proops, Oxford University Press (2021). © Ian Proops.
DOI: 10.1093/oso/9780199656042.003.0002

1.2 Illusions: logical, empirical, and transcendental

Kant characterizes illusion in general as the enticement [*Verleitung*] to erroneous judgment (A 293/B 350). It comes in logical, empirical, and transcendental flavours. *Logical* illusion is the enticement to logically fallacious inference. It is operative when, owing to inattention to the logical rule, one adjudges an argument that is in fact invalid to be valid (A 297/B 353). This variety of illusion counts as 'artificial', in Kant's view, because there is nothing in our make-up dictating that we must ever experience it. Because logical illusion arises *merely* from inadvertence (A 298/B 354), it 'entirely disappears' as soon as we attend to the correct logical rule—presumably with the knowledge that it *is* correct (A 297/B 353).[2] Insofar as it possesses a certain fleetingness and fragility, logical illusion contrasts sharply with transcendental illusion, which enjoys a certain robustness befitting its character as 'natural' and 'unavoidable' (A 297–8/B 353–4).

Empirical illusion comprises instances of naturally occurring perceptual illusion—and paradigmatically, naturally occurring optical illusions (A 295/B 351–2). A subject in the grip of an empirical illusion—that is, a subject who is perceptually appeared to in a certain deceptive way but who is not hallucinating—is tempted to make a false judgment about an actually perceived object. And when the subject succumbs to this temptation their faculty of judgment is 'misled through the influence of the imagination' (ibid.). Kant mentions as examples of this kind of illusion the appearance that the sea is higher in the middle than at the shores, as well as the so-called 'moon illusion', in which the moon appears larger on the horizon than at its zenith (A 297/B 354).

Officially, the source of transcendental illusion is the *understanding* (A 581/B 609), but unofficially it is *reason*. Kant, after all, characterizes the dialectical inferences of pure reason as having sprung from 'the nature of reason' (see, for example, A 339/B 397) and he asserts that the illusion involved in the paralogisms has its *ground* in the nature of human reason (A 341/B 399). The discrepancy between the official and unofficial accounts can be explained by the fact that he sometimes treats the understanding as a sub-faculty of the faculty of reason (in the broad sense). For example, in the dynamical antinomies, when he portrays reason as seeming to be in conflict with itself, he treats the antithesis as satisfying the needs of the understanding, the thesis those of reason (now in the narrow sense) (see A 531/B 559).

Kant sees transcendental illusion as sharing three important affinities with empirical illusion. First, as we have already mentioned, it is *natural* and, since

[2] Kant understands 'applied logic' as the discipline that deals with (among other things) attention, including, one presumes, attention to the logical rule (A 54/B 79). In applied logic, one learns (among other things) how to dispel logical illusion.

incapable of being dispelled, also *permanent* (A 298/B 354).[3] Just as anyone equipped with properly functioning visual faculties will, Kant supposes, be subject to the moon illusion, so anyone equipped with a properly functioning faculty of reason will be subject to transcendental illusion.[4]

The second feature that Kant takes transcendental illusion to share with empirical illusion is its status as 'not irresoluble' [*nicht unauflöslich*] (A 341/B 399). He means that although it can never be dispelled, transcendental illusion remains in principle ultimately innocuous insofar as its usual consequences are *correctible* (A 644/B 672).[5] Thus, just as one need not be 'taken in' by an optical illusion—in the sense of actually forming a judgment or belief on its basis—so a subject experiencing transcendental illusion need not be led to adopt any dogmatic belief on its basis. We can, as Kant says, 'forestall the error' (A 339/B 397; compare A 644–5/B 672–3), even if 'we can never fully rid ourselves of the illusion' (A 339/B 397).[6]

It would be hard to exaggerate the importance Kant attaches to the distinction between avoidable error and inevitable illusion. As Michelle Grier has observed, if one fails to draw it, Kant will seem to be making the patently false claim that the unsound arguments of the Transcendental Dialectic are themselves inevitable.[7] With the distinction drawn, however, it becomes clear that rather than vainly cautioning us against inevitable error, Kant is merely alerting us to the existence and workings of transcendental illusion in the hope of fortifying us against its

[3] That Kant regards transcendental illusion as permanent is clear from his remark that transcendental dialectic (that is, the critique of illusion) 'can never bring it about that transcendental illusion should [not merely cease to deceive us] but also disappear (as logical illusion does [when detected]) and cease to be an illusion' (A 297–8/B 354).

[4] Kant seems to imply that the optical illusions he describes affect all (properly functioning) human beings. But whether that is in fact so is a question for empirical psychology.

[5] I avoid saying that the illusion *itself* is 'correctible' because this might suggest a view Kant did *not* hold, namely, that transcendental illusion is capable of being neutralized by a compensating counter-illusion. As Roy Sorensen has remarked (in conversation), it is puzzling that Kant held out no hope of someone's eventually designing such a counter-illusion (as British road engineers seeking 'traffic calming' nudges have designed roadway striping to create the illusion of speed). Why was he pessimistic about finding such a permanent fix to the problem of transcendental illusion? Wouldn't that have been better than simply recommending *Critique* as an endlessly needed—and difficult to apply—expedient? This is a good question. Kant was not, after all, ignorant of the idea of a compensating counter-illusion. He knew, for example, about John Dollond's 'achromatic doublet'—a pairing of a concave flint glass lens with a convex crown glass lens to correct for chromatic aberration in telescopes (29: 915). Of course, he would also presumably have known about corrective spectacles, and very possibly also—given his lively interest in architecture—about the illusion-compensation theory of column entasis espoused by Vitruvius (see *The Ten Books on Architecture*, III, iii, 13). Incidentally, for similar reasons, I avoid describing the illusion as 'resistible'—as indeed does Kant, who in fact explicitly calls it 'irresistible' (A 642/B 670).

[6] Accordingly, when at A 642/B 676 Kant says that the illusion's deception is something that one can 'barely' [*kaum*] prevent through the most acute critique, he does not mean that one *cannot* prevent it but merely that one can *only just* prevent it.

[7] See Michelle Grier, 'Illusion and Fallacy in Kant's First Paralogism', *Kant-Studien*, 83 (1993), 257–82 at 263–4 and also her book *Kant's Doctrine of Transcendental Illusion* (Cambridge: Cambridge University Press, 2001) at 116 and 28–30.

customary effects. In particular, by attending to the existence of transcendental illusion, we will, he hopes, be better able to resist the appearance of soundness attaching to certain arguments within speculative dogmatic metaphysics that are in fact not known to be sound. His idea is that, being able to resist forming the judgments that constitute the conclusions of these arguments, we will be better equipped to cast a sceptical eye over the arguments themselves.

The third feature that transcendental illusion shares with empirical illusion relates to an aspect of its independence from belief and judgment, namely, its *persistence* in the face of countervailing beliefs. Transcendental illusion does not 'cease even though it is uncovered and its nullity [*Nichtigkeit*] clearly seen into by transcendental criticism' (A 297/B 353). In other words, it survives detection and even—Kant is perhaps saying—*explanation*. In today's jargon we would say that, like the Müller-Lyer illusion, transcendental illusion 'resists cognitive penetration'.[8] This means that a subject who, being wise to the illusion, forms a belief contrary to the erroneous one it standardly tempts a person to form will still experience the illusion and still feel tempted to form the false belief.

The persistence of transcendental illusion has implications for the practice of critique. Owing to it, Kant thinks, human reason will be perpetually propelled into 'momentary aberrations that always need to be removed' (A 298/B 355). And in consequence the work of transcendental criticism will never be finished.

1.3 The sources of transcendental illusion

Although he believes that transcendental illusion has a number of 'sources' (A 581/B 609) or 'causes' (A703/B 731), Kant singles out one cause as pre-eminent. This, he tells us, is the fact that

> in our reason (considered subjectively as a human faculty of cognition) there lie fundamental rules and maxims for its use, which [while subjective in character] have entirely the look of objective principles, and through them it comes about that the subjective necessity of a certain connection of our concepts for the benefit of the understanding is taken for an objective necessity, the determination of things themselves. (A 297/B 353)

Kant characterizes one of these subjective rules or maxims as a 'demand of reason'. 'Multiplicity [*Mannigfaltigkeit*] of rules and unity of principles', he says, 'is a demand of reason, [something demanded] in order to bring the understanding into thoroughgoing accordance with itself' (A 305/B 362). This demand is a

[8] See A 581/B 609; and compare A 703/B 731 and 7: 149–50.

methodological principle; in Kant's language it is a 'subjective' principle (or law) or, again, a 'logical prescription [*Vorschrift*]' (A 309/B 365). Kant is telling us that the use of our reason in theory construction is governed by the norm that we should strive to bring our knowledge into a system in which the richest variety of rules is unified by the smallest—or, better, most explanatorily deep—set of principles, something, he goes on to make clear, that involves, among other things, minimizing the theory's conceptual inventory (I take it that 'principles' therefore are fundamental components of a theory—basic concepts as well as basic propositions).[9] This last point is not spelled out in the text, but it is suggested by Kant's referring to the demand in question as a 'subjective law of housekeeping [*Haushaltung*]' aimed at economizing on the 'supplies of the understanding' (A 306/B 362). Kant, I take it, is deploying this image of domestic frugality to suggest that one aspect of theoretical unity is the minimization of the number of concepts that the theory treats as undefined—the minimization, in Quinean terms, of the theory's 'ideology'. The demand in question, Kant is suggesting, is *inter alia* the demand to minimize ideology through conceptual reductions in which one concept is characterized in terms of others. That this is part of what Kant has in mind is suggested by his speaking of effecting the reduction in the number of concepts by *comparing* one with another (ibid.); for the language of comparison is the language he employs when speaking of reduction in other contexts. (See, for example, A 649/B 677, where he envisages the reduction of one power to another.) Kant's image of efficient housekeeping, I take it, is intended to work by suggesting that just as a housekeeper stands under a *ceteris paribus* prudential obligation to use up the household's provisions (so that food doesn't spoil), so the empirical scientist stands under a *ceteris paribus* prudential obligation not to leave concepts unreduced (so that theories don't conceptually bloat). If this is right, then Kant's implied image of an over-stocked pantry or storehouse is playing a role in connection with ideology exactly parallel to the role played by Quine's image of an over-populated slum in connection with ontology.[10]

All of this suggests that we may formulate the aforementioned demand of reason as a norm on theory construction that tells us to seek that theory among those fitting the data which maximizes the combination of explanatory power and ideological parsimony. Kant doesn't consider whether the principle in question could always identify a uniquely preferred empirically adequate theory, but since trade-offs between explanatory power and parsimony are plausibly possible, this question is one he might have considered.

[9] The qualification about explanatory depth is needed because a theory with a single, highly disjunctive or gerrymandered axiom would not count as highly unified in the sense Kant clearly intends.

[10] The image of concepts as residing in 'the storehouse of the mind' would have been familiar to Kant from Descartes's fifth *Meditation*.

Kant's model for the procedure by which one arrives at reason's methodological principles or rules is 'reason's formal and logical procedure in syllogisms' (A 306/B 363). This is the procedure in which one constructs polysyllogisms, by seeking ever higher logical grounds or logical conditions for a given conclusion.[11] In the *Logik Hechsel* Kant illustrates this procedure with the following example: 'Everything that thinks is simple, the soul thinks, hence it is simple. Everything that is simple is indivisible. The soul is simple, hence it is indivisible.'[12] Here the first syllogism purports to establish the minor premise of the second (namely, 'The soul is simple'). To supply the first syllogism when presented with the second is just to take the first step in a procedure which involves linking together ever higher syllogisms into a polysyllogistic chain, advancing to ever more general principles (considered as premises) as one does so.

Although this particular illustrative example fails to do justice to the idea of ideological parsimony, it is adequate for Kant's purposes—namely, to illustrate how a merely methodological principle (a 'logical maxim') might, through a misunderstanding, be taken for a substantive principle generating dogmatic metaphysics (a 'principle of pure reason') (A 306–7/B 363–4).[13] Kant likes the example partly because it makes vivid the point that sometimes we seek unification at a level of abstraction higher than that of the subject matter of our science. His thought is that just as we do not need to intuit particular features of objects in order to apply the norm of seeking to unify our judgments about them by searching for ever more general logical conditions in constructing polysyllogisms, so we do not need to note any particular necessary features of objects—such as that every occurrence has a cause—in order to apply the general norm on theory construction in theorizing about them (see A 306/B 363). The example of constructing polysyllogisms thus serves to illustrate the idea, which Kant is clearly anxious to underscore, that the theoretical unity we are urged to seek by his general norm on theory construction differs in *kind* from the object-level unity—expressed, for example, by the causal principle—of possible experience (A 307/B 363). One imagines that he feels the need to stress this point as forcefully as he does because he thinks that, owing to transcendental illusion, we are especially prone to overlook the distinction that it makes salient.

[11] See, A 307–8/B 363–5 and note Kant's talk at A 500/B 528 of 'the logical demand to assume complete premises for a given conclusion'. For an illuminating treatment of Kant's account of conditions in connection with polysyllogisms, see Tobias Rosefeldt, 'Subjects of Kant's First Paralogism', in A. Stephenson and A. Gomes (eds), *Kant and the Philosophy of Mind: Perception, Reason, and the Self* (Oxford: Oxford University Press, 2017), 221–44.

[12] Tillmann Pinder, *Immanuel Kant Logik-Vorlesung Unveröffentlichte Nachschriften, Logik Hechsel; Warschauer Logik* (2; Hamburg: Meiner, 1998) at 484.

[13] That one use of 'logical' as qualifying 'principle' or 'maxim' in the first *Critique* has the meaning of 'methodological' (or 'subjective') is apparent in many places, but perhaps most clearly of all at the end of the first full paragraph of A 648/B 676. This same passage makes clear that 'logical' in this sense is opposed to 'transcendental' and 'objective'.

Kant's illustrative example is designed also to bring out how *repeated* application of his envisaged general norm on theory construction can, through its capacity to generate an ascending polysyllogistic chain, lead to the construction of a *regress* of ever more general conditions (A 307/B 364). And it further helps to bring out that there is nothing *in itself* problematic about applying the methodological norm in question. Trouble arises, Kant thinks, only because we are inclined to mistake reason's injunction to proceed in our theorizing by seeking ever higher logical conditions—whether premises or principles—for a *doctrine* of reason, namely, the thesis that we inhabit a world characterized by a series of world-states or objects of sense ordered by the *real* conditioning relation and either containing an unconditioned condition as its first member or, being infinite, containing an infinite series of conditions that is itself unconditioned. Either way, Kant thinks, when we yield to this inclination we lapse into error.[14]

Since the example of polysyllogism construction fails to illustrate the part of the general norm that urges us to advance theoretical unification by reducing one concept to another (or others), it is only a partial illustration of the procedure recommended by the norm on theory construction Kant is envisaging at A 305/B 361–2. This is not a serious problem, however, for two reasons. First, as Kant himself stresses, the example of polysyllogism construction by itself already suffices to illustrate the two points just mentioned—for *these* purposes, there is no need to advert to the second part of the general norm. Second, another discussion in the first *Critique* provides the material for an especially clear-cut illustration of how misinterpreting the norm of maximal ideological parsimony might lead us into dogmatism. This example involves the *ontological* parsimony of powers rather than an ideological parsimony, but it is readily adaptable to our purposes. It concerns, in particular, the generation of the rational psychologist's belief that all the powers of the soul—imagination, memory, wit, and so forth—are reducible to a single fundamental power (A 649/B 677).[15] Kant's idea is that by

[14] Kant, it should be noted, has no *general* objection to an infinite series' bearing the real conditioning relation to something. He allows, for example, that a given moment of time is (really) conditioned by the infinite series of times preceding it. He says: 'The entire elapsed time, *as condition of the given instant*, is thought necessarily as given' (A 412/B 439, emphasis added). He is able to allow that a given instant of time is borne the relation of real conditioning by an infinite series of times because time is an infinite *totum analyticum*. What he does reject (as leading to antinomies) is the assumption that real conditioning relation should be borne to something by an infinite *totum syntheticum* (for example, by the series of world-states prior to a given world state). We discuss the notions of *totum analyticum* and *totum syntheticum* in Chapter 9. Note, incidentally, that Kant does not say—and, I think, does not hold—that an instant is conditioned by any allegedly 'immediately previous' time. His view would seem to be that no time is immediately preceded by another (see B 413–14, for Kant's apparent rejection of the view that there could be two distinct times with no time between them).

[15] This isn't, however, quite how every rational psychologist would *express* the view Kant has in mind. Baumgarten, for example, regards the faculties that are derivable from the soul's basic power not as powers properly speaking, but rather merely as modes of the power of representation. See Baumgarten, *Metaphysics* §744; and Gary Hatfield, 'Baumgarten, Wolff, Descartes, and the Origins of Psychology', in Courtney D. Fugate and John Hymers (eds), *Baumgarten and Kant on Metaphysics* (Oxford: Oxford University Press, 2018), 61–77 at 66.

confusing the demand for ontological parsimony in our theories with imagined knowledge of unification in the phenomena themselves we are led to suppose, without justification, that there is a single power of which all the others are determinations or sub-kinds (sub-powers). As Kant would have been aware, Wolff believed that the soul did indeed possess only a single power, namely, its power of representing the world from the point of view of its associated body (see Chapter 2). In Wolff, therefore, Kant has a flesh-and-blood example of a philosopher to whom the present diagnosis might apply.

Although this example involves an ontological rather than conceptual reduction, it may be adapted to a yield an illustration of a case in which a legitimate push for *conceptual* reduction leads to a parallel mistake. One might, for example, try to *define* memory as, say, the power to veridically *represent* something as past, thereby seeking to diminish the number of concepts needed to describe the phenomena. Such a conceptual reduction might be a reasonable thing to attempt, but in Kant's view such a project will tend to encourage the unfounded belief that the relevant powers themselves admit of a corresponding ontological reduction.

Noting the centrality of Kant's idea that we are inclined to confuse a methodological, prescriptive principle with a constitutive one, Michelle Grier has recommended that we interpret Kant as taking transcendental illusion to be 'manifested' by the conflation of two principles. The idea is that we conflate a 'logical prescription' governing theory construction (A 309/B 365), namely, *P* (Grier's 'P1'):

> *P*: Find for the conditioned cognitions of the understanding the unconditioned whereby its [that is, the understanding's] unity is completed (cf. A 307/B 364)

with a descriptive claim about the world, namely, *D* (Grier's 'P2'):

> *D*: If the conditioned is given, the whole series of conditions subordinated to one another, a series which is therefore itself unconditioned, is likewise given (cf. A 307–8/B 364 & B 436).[16]

I agree with Grier that this conflation—mistaking *P* for *D*—is certainly included among the mistakes encouraged by transcendental illusion, but I think that

[16] See Grier, 'Illusion and Fallacy'. And also Grier, *Kant's Doctrine* at 268–79. My labels '*P*' (for 'prescription') and '*D*' (for 'description') are intended as arguably more suggestive replacements for Grier's 'P1' and 'P2'. Although I agree with much of what she says, I disagree with Grier on three points of detail. First, I would not endorse her characterization of *D* as something we must assume as a condition of 'using' *P* (Grier, *Kant's Doctrine*, 126). Being a prescription, *P* is not something that can be used, but only something that can be complied with or contravened. Instead, I think the correct thing to say about the relationship between *P* and *D* is that the illusion that *D* is true is one that is, in Kant's view, inevitably generated in rational human minds in virtue of their being subject to the demand of reason expressed by the prescription *P*. Second, as we will see in Chapter 5, it is in fact not quite *D* itself but a closely related claim that expresses the content of the illusion. Third, I differ from Grier on the form of the first paralogism (see Chapter 4).

transcendental illusion in fact admits of a more general characterization. Most generally characterized, it is the illusion that something that is in fact (in a broad sense) merely subjective—methodological demands on theory construction included—is an objective feature of the world.[17] Nonetheless, the mistake of taking *P* for *D* is certainly one central *instance* of the error that can arise when we succumb to this illusion—and, arguably, it is the most important such instance.

Prescription, *P*, enjoins us to seek the most unified theory that is empirically adequate. Descriptive claim, *D*, on the other hand, declares that for every phenomenon (in the theory-neutral sense of that term) there exists a series constituted by its unconditioned condition. In Kant's view, when through a 'misunderstanding' we are deceived by transcendental illusion, the 'need of reason' expressed by *P* is 'taken for a transcendental principle of reason' that 'over hastily postulates . . . an unlimited completeness in the series of conditions' (cf. A 309/B 366).[18] In other words, we mistake the prescription, *P* for the putative statement of fact, *D*. In Kant's jargon, we mistake a 'regulative' for a 'constitutive' principle.[19]

Although Kant characterizes the 'misunderstanding' (A 309/B 365) that leads to certain fallacious 'inferences of reason' in terms of our having 'taken' the need of reason expressed by the prescription, *P*, for the truth of the principle, *D* (A 309/B 366),[20] the path from the former to the latter may, I think, be usefully be characterized in more fine-grained terms. It seems that our first mistake must be to misconstrue a prescription concerning *cognitions* (that is, in this context, rules and maxims governing theory construction) as a prescription concerning the worldly objects themselves. That is to say, we hear *P* as saying:

*P**: Find for the object of any given cognition, that object's unconditioned series of conditions.

It seems plausible that, being a command of *reason*, the legitimate prescription, *P*, will be viewed as *authoritative*. Moreover, *P* is plausibly experienced as *reasonable* because reason's authority strikes those subject to it as reasonably exercised. But since the prescription, *P*, strikes us as reasonable and authoritative, the same will be true of our misconstrual of it, namely, *P**. We will accordingly find it natural to

[17] This broad characterization of transcendental illusion is bound up with yet another understanding of dialectic, not now as the critique of illusion, but rather as 'the doctrine of the subjective laws of the understanding insofar as they are taken for objective' (R 1579; 16: 23).

[18] Kant gives this explanation in the form of a lengthy question in the course of announcing the topic to be investigated. We might paraphrase the question as: 'Has the need of reason expressed by *P* been taken for the transcendental principle expressed by *D*?' It is clear from his subsequent discussion that, as far as traditional dogmatic metaphysics goes, Kant's answer is yes.

[19] See A 619–20/B 647–8; A 647/B 675; A 686/B 713; A 690/B 718; and *Prolegomena*, §56, 4: 350.

[20] Kant also implies that *instances* of *P* are apt to be taken for instances of *D*, for he says that the various specific fundamental maxims and rules contributed by human reason and governing its use 'look entirely like objective principles' (A 297/B 313).

suppose that what the apparently authoritative prescription P^* enjoins us to seek is in fact actually there to be found. In this way, the appearance of the standing authority of P (and so also of P^*) will generate in us the permanent inclination to regard D as true.

Or so, at least, runs a plausible interpretation of Kant's first-pass view. However, I will argue in Chapter 5 that his considered view is that it is not D itself that states the content of this sub-illusion, but rather the related proposition, D^*, which omits the reference to a series: 'If a conditioned object is given, an absolutely unconditioned object is likewise given.' Because the reasons for this qualification are complicated, however, it will be convenient, for the time being, to continue our preliminary exposition with reference to D.

In portraying the immediate illusion lurking behind much of dogmatic speculative metaphysics as the illusion that D is true, Kant may seem to be lumping his predecessors together in an unduly high-handed way. Is it really true that much of dogmatic speculative metaphysics flows from a single principle? Such qualms may, however, be alleviated somewhat by the observation of Predrag Cicovacki that D can be regarded as one formulation of the Principle of Sufficient Reason (hereafter '*PSR*')—a principle which does indeed have a claim to be a substantive principle which generates many of the claims of dogmatic speculative metaphysics.[21] This idea might also explain why Kant refers to D as 'this supreme principle of pure reason' (A 308/B 365).

In Kant's view, we are inclined to regard D not just as true but as *necessarily* true. This circumstance, he maintains, is owed to our readiness to mistake the 'subjective necessity' of 'a certain connection of our concepts' for an 'objective necessity, the determination of things in themselves' (A 297/B 353). Kant doesn't say what this 'subjective necessity' is supposed to consist in; but it is plausibly the necessity involved in the following conditional injunction: 'If you wish to proceed rationally in inquiry, you *must* seek, for the object of any cognition, the series constituting its unconditioned condition.' The necessity involved here counts as 'subjective' because the imperative is conditional upon our having a certain goal or desire.[22] Kant supposes that, owing to transcendental illusion, this subjective necessity is apt to be misconstrued as the objective necessity involved in the

[21] Predrag Cicovacki, 'Kant's Debt to Leibniz', in Bird Graham (ed.), *A Companion to Kant* (Oxford: Blackwell 2006), 79–92 at 87. See also James Kreines, 'Metaphysics Without Pre-Critical Monism', *The Bulletin of the Hegel Society of Great Britain*, 29 (1–2) (2008), 48–70 at 49; and Omri Boehm, *Kant's Critique of Spinoza* (Oxford: Oxford University Press, 2014) at 51. Two caveats: first, D is plausibly equivalent to the *PSR* only on the Kantian assumption—to be discussed in Chapter 15—that nothing can be a condition of itself. If something could be its own condition, then the demands of the *PSR* might be satisfied while D was false. Second, I disagree with Boehm's view that P (his 'P1') is a formulation of the *PSR* (Boehm, *Kant's Critique of Spinoza*, 51). It is rather a methodological principle urging us to pursue economically constructed theories. I had overlooked this last point in Ian Proops, 'Kant's First Paralogism', *Philosophical Review*, 119 (4) (2010), 449–95.

[22] Elsewhere, Kant indicates that he sees the prescriptive force of a certain postulate as similarly conditional on our desires. We should, he says, think of events in such a way that we can apply our

claim: 'Necessarily, for the object of any cognition, the series constituting its unconditioned condition exists.' The illusion that *D* holds *of necessity*, then, is itself a component of transcendental illusion.

To clarify the slide that I take Kant to be envisaging it will be helpful to break it up into three transitions between four claims:

[1] If you desire to proceed rationally in inquiry, you must seek—that is, strive to formulate—the most unified theory, where unification is understood in terms of finding ever higher logical conditions (that is, principles) for one's cognitions.
[2] If you desire to proceed rationally in inquiry, you must seek, for the object of any cognition, the series constituting its unconditioned (real) condition.
[3] You must seek, for the object of any cognition, the series constituting its unconditioned (real) condition.
[4] Necessarily, for the object of any cognition, the series constituting its unconditioned (real) condition exists.

Claim [4] is just the claim that principle *D* (slightly reformulated) holds with necessity.

As I read him, Kant supposes that we are apt to confuse [1] with [2] because a certain principle governing theory construction looks like a principle guiding worldly investigation. This *seeming*, I take it, is an instance of transcendental illusion because it involves seeing a subjective unity (the unity that would be enjoyed by a regressing chain of syllogisms if it could be extended to a logically unconditioned minor premise) as an objective unity (the unity that would be enjoyed by a series of worldly states if it were to have an ultimate unconditioned real condition).[23] Thus, when we are taken in by the sub-illusion that [1] says what is said by [2] we are apt to slide from acceptance of [1] to acceptance of [2]. The confusion involved in the slide from [1] to [2] is thus a confusion about the nature of the object for which the inquirer is seeking.

Beyond this, we are further inclined to misconstrue [2] as meaning [3] simply because, being rational beings, we *do* naturally desire to proceed rationally when we inquire. Since this desire is so natural and pervasive its presence is easily overlooked, and we thus fail to appreciate that the necessity in [3] is merely the conditional necessity expressed in the consequent of [2]. The sub-illusion that [2] says the same as [3] is the illusion that an intellectual—as opposed to practical—*hypothetical* imperative is an intellectual *categorical* imperative.

understanding to them 'if we want our understanding to be in unison with itself in accordance with principles' (R 6109, 1783–84, 18: 457).

[23] Recall that, as Kant sets things up, at each step of the polysyllogism it is the *minor* premise that is derived by a further syllogism.

Finally, we are further inclined to misconstrue [3] as saying [4] because we hear [3] as specifically an injunction of *reason*, and so regard it as rational. This in turn inclines us to believe that what we are enjoined to seek actually exists. We are thus inclined to misconstrue the injunctive modality in [3] as the alethic modality in [4]. That is to say, we are inclined to mistake the 'must' of (what we regard as) a binding intellectual categorical imperative for the 'must' of an indicative claim purporting to express a necessary fact.[24]

Each of these sub-illusions may be regarded as a matter of our being inclined to take something that is, in some respect or other, merely subjective for something objective. Each one, therefore, is plausibly viewed as an instance of the general inclination Kant identifies sometimes with transcendental illusion and sometimes with its cause. The upshot is that we are inclined to slide all the way from an acceptance of the (eminently reasonable) hypothetical imperative of inquiry [1] to an endorsement of the strong metaphysical claim [4]. But whether we do so in practice will depend on whether we have been tipped off by the critical philosophy to the instances of transcendental illusion that constitute the misleading appearance that the content of [1] is the same as that of [2], that the content of [2] is the same as that of [3], and so forth. Transcendental illusion is thus to be invoked in *explaining* the slide between judgments, but an aspect of it can also be thought of as the *product* of the slide insofar as the three sub-illusions combine to transform the veridical appearance that [1] is a maxim of rational inquiry into the illusory appearance that [4] is true.

In support of our explanation of the steps leading from [2] to [3] and from [3] to [4], we may observe that on one occasion Kant himself draws the distinction between objective and subjective necessity in precisely the terms we have suggested—although he does so in a less familiar setting. In religion lectures from the mid-1780s, speaking of his earlier attempted proof of the existence of an *ens realissimum* (or 'most real being') in his 1763 work *The Only Possible Basis for a Demonstration of the Existence of God* (hereafter '*OPB*'), he says that although this proof is, contrary to his former hope, unable to establish the objective necessity of an original being, it nonetheless succeeds in establishing the subjective necessity of accepting such a being (28: 1034).[25] His point in these lectures is that although I cannot demonstrate the existence of a being that is the ground of everything possible, my reason makes it necessary for me to *accept* the existence of such a being because otherwise I would be unable to *cognize* in what in general the possibility of something consists (ibid.).[26] Kant thus comes to believe that his earlier 'proof' in *OPB*, although it hadn't attained its professed goal, had

[24] I am indebted to Derek Haderlie for help in formulating this last claim.

[25] *OPB* is dated 1763, but was in fact published in 1762.

[26] This passage has been raised to prominence by Andrew Chignell. See 'Belief in Kant', *Philosophical Review*, 116 (3) (2007), 323–60 at 349. I follow his translation.

nonetheless succeeded in establishing the merely subjective necessity involved in the conditional claim: 'If we want to cognize in what in general the possibility of a thing consists, then we must accept that the *ens realissimum* exists.' Because in *OPB* Kant had taken the 'proof' to establish the dramatically stronger claim that the *ens realissimum* exists, the clear implication of his later remarks is that he believes that we are inclined in this case to slide from a hypothetical imperative of inquiry, expressed as a conditional, to an objective claim of necessity—namely, the claim that the *ens realissimum* must exist—precisely because we *do* wish to cognize in what in general the possibility of a thing consists. Kant is appealing to this idea in his religion lectures in order to diagnose his own earlier dogmatic mistake.

I think a similar slide is involved in the transition from [2] to [4]. To be sure, in the case of our hypothetical imperative, [2], what figures in the consequent is not, as in *OPB*, the notion of *accepting* a proposition. It is rather that of *seeking* the unconditioned. This difference of attitude, however, does nothing to undermine the parallel between the relevant passages with respect to Kant's use of the terms 'subjective' and 'objective' in their role as qualifying the term 'necessity'.

Our description of the process leading from [1] to [4] amounts to an account of how rational human agents can be led to form certain *beliefs* within dogmatic speculative metaphysics—beliefs, namely, in instances of the general claim, [4]. But more would need to be said if Kant is to invoke transcendental illusion—as he seems to wish to do—to explain why philosophers of the calibre of a Descartes or a Leibniz should have come to believe that they have genuine *knowledge* of instances of [4].[27] Kant leaves this lacuna unfilled, but we might fill it by continuing our story in the following way.

Having arrived at a given dogmatic speculative belief—that is, a given instance of [4]—Kant's dogmatic speculative metaphysician, not supposing that his belief could have any source other than pure reason, proceeds to imagine that his grounds for it must be principles known a priori. He then tries to make explicit the reasoning he takes himself to have implicitly relied on in forming his belief. This involves constructing a proof of the relevant instance of [4]. Since he imagines this proof to be sound and to proceed by a priori reasoning from

[27] 'Through the explanation of illusion', Kant says, 'one grants to the one who erred a kind of fairness' (9: 56). He means, I take it, that by explaining the illusion prompting an error we render that error comprehensible as arising from something other than obtuseness. Nowhere, it seems, is the need for such an explanation more urgent than in connection with Leibniz. For Kant's estimation of Leibniz's intellect could hardly be higher. 'No philosopher', he rhapsodizes, 'has ever shown such extensive skill in philosophizing dogmatically as Leibniz' (24: 804). Charity, then, is one reason to posit transcendental illusion. Another, I think (though this is more speculative), is the need to posit a *real* ground for error. For, unlike Descartes, Kant sees error [*Irrtum*] as a reality rather than a lack [*Mangel*] or negation (see 28: 1272). It would therefore have been natural for him to seek a reality to ground this reality. Transcendental illusion would, I think, have filled this bill.

principles known a priori, he thereby comes to believe himself possessed of theoretically based *knowledge* of the metaphysical proposition in question.

1.4 The simplified account

Our story has emphasized passages in which Kant suggests that the first error we make upon succumbing to transcendental illusion is that of confusing a logical (that is to say, methodological) imperatival principle—a conditional injunction concerning theory construction—with a real or metaphysical imperatival principle—a conditional injunction concerning the discovery of worldly phenomena. However, our account would be incomplete if we did not give due weight to the existence in the *B* edition of a simpler account of how transcendental illusion operates. This rival account makes no mention of the confusion of a logical with a real principle. Instead, Kant puts the blame on a misapplication of reason's legitimate demand for the unconditioned. In the *B*-edition preface, he says:

> [W]hat necessarily impels us to go beyond the boundary of experience and of all appearances is the **unconditioned**, which reason demands in things in themselves; [within the realm of things in themselves] reason necessarily and with full right—demands this unconditioned for everything conditioned, thus demanding that the series of conditions be completed by means of that unconditioned. (B *xx*)[28]

One clear implication of this remark is that so long as the 'things' in question are things in themselves, the categorial injunction to seek the unconditioned in things (rather than in theories) is perfectly legitimate. Kant says nothing here to explain why this demand should cause us to illegitimately go beyond the boundary of experience, but the following story seems likely. If, having succumbed to transcendental illusion, we mistake the objects of sense for things in themselves, we will go beyond experience in attempting to comply with this (in itself legitimate) demand of reason. For, taking the objects of sense to be things in themselves, we will take an injunction that, in truth, applies only to the latter, to apply to the former. We will thus suppose that what it enjoins us to seek is in fact there to be found. And so we will come to believe that the spatio-temporal world series contains an unconditioned object or state of affairs (or, alternatively, a series of them) that is not an object of experience.

[28] And compare the following remark from a reflection: 'The proposition that if the conditioned is given, the whole series of all conditions through which the conditioned is determined is also given is, if I abstract from the objects or take it merely intellectually, correct' (R 5553 (18: 223); compare 4: 354).

This new account dispenses with steps [1] and [2] of our reconstruction, rejoining it at step [3]. It seems to have been an attempt by Kant to streamline his explanation of how transcendental illusion induces error. It does so by emphasizing that the content of transcendental illusion is equivalent to Transcendental Realism. Unfortunately, however, Kant fails to remove certain traces of his more complicated (and presumably earlier) account from the second edition. This is not all that surprising because those traces occur in the parts of the *Critique* that he did not substantially revise.

1.5 The necessity of transcendental illusion

As we have seen, Kant believes that rational human beings *inevitably* experience transcendental illusion.[29] Unsurprisingly, commentators have wondered what his argument for this inevitability thesis is supposed to be.[30] Possibly, Kant intends to be offering an argument by analogy with the perceptual case. The idea would be that just as human beings inevitably experience various perceptual illusions in virtue of possessing properly functioning perceptual faculties, so they inevitably experience transcendental illusion in virtue of having a properly functioning rational faculty. The grounds for such an inference would include the numerous similarities noted between (theoretical) reason and sense perception: each faculty, for example, operates on an input of a certain characteristic kind (in the case of perception, sensation; in the case of reason, judgment), each is apt, in virtue of its very organization, sometimes to lead the judgmental faculty astray, and so forth. The inference would consist in affirming one more similarity on the basis of these: namely, that whatever it is that inclines the judgmental faculty towards error must, like perceptual illusion, be universal in rational human subjects.

Such an argument belongs to a genre of which Kant certainly approved; for as we will see in later chapters, he recognizes arguments by analogy as legitimate so long as they are held to be capable of producing only 'empirical certainty' (9: 132–3). The interpretive proposal, then, has its merits. But it also faces a

[29] Kant associates transcendental illusion with specifically *human* reason. He does not seem to think that angels would be subject to it, and he plainly does not regard non-human animals as subject to it. If the first point sets him in tension with the tradition (for wasn't Lucifer tempted?), the second sets him against experience. For Kant's most general characterization of transcendental illusion—as the systematic appearance that something in fact subjective is an objective feature of the world—would seem to imply that some non-human animals are subject to it. Take, for example, my cat, Gremlin. She sometimes suffers from UTIs. When thus afflicted, she will not urinate in the same place twice. She behaves as if she is taken in by the illusion that something subjective (her pain) is an objective (and localized) feature of the world, as if there existed painful areas of the litter box (or sofa!) which, once discovered, are to be studiously avoided.

[30] For a discussion of relevant literature see Michael Rohlf, 'The Ideas of Pure Reason', in Paul Guyer (ed.), *The Cambridge Companion to Kant's Critique of Pure Reason* (Cambridge: Cambridge University Press, 2010), 190–209.

difficulty—albeit one of which Kant seems unlikely to have been aware—the difficulty, namely, that its motivating parallel is in fact undermined by the empirical data. Unfortunately for Kant, certain optical illusions turn out *not* to be universal in human beings, and so cannot be psychologically necessary. For example, subjects brought up in different environments within a single culture have been found to differ in their susceptibility to the Müller-Lyer illusion.[31] One premise of the argument by analogy therefore fails.

A second possible argument for the necessity thesis would appeal to the idea that reason has an essence. It is of the essence of reason—or so this argument would maintain—that rational beings feel motivated by an injunction of reason to ask theoretical *why*-questions and to press them as far as possible.[32] To engage in such inquiries, this line further holds, is therefore a fundamental need of human reason. From this assumption one would conclude that the production of transcendental illusion is, for this very reason, inevitable in rational human beings, for it is inevitably produced by the imperatives governing this kind of inquiry. This strategy has the advantage of making use of the idea of a 'need of reason'—a notion to which Kant himself often appeals. It has the obvious weakness, however, of being anthropologically dubious. For it is not at all clear that human rationality must always lead to scientific or metaphysical curiosity. In some individuals it might rather find expression in, for example, one or more of the following activities: chess, music appreciation, pigeon keeping, cineastry, cooking, carpentry, or floral art.

For these reasons, I'm inclined to doubt that Kant has a strong argument for the thesis that transcendental illusion is necessary—even psychologically—in rational human beings. But perhaps this matters less than it appears. After all, what should matter for Kant—or so one might have thought—is merely to draw attention to the existence of transcendental illusion and to make a plausible case that it is sufficiently pervasive and enduring as to be capable of explaining the *apparently* universal and unceasing human drive to engage in dogmatic speculative metaphysics. In defending such a claim Kant would, I think, be on firmer ground.

[31] A. Ahluwalia, 'An Intra-Cultural Investigation of Susceptibility to "Perspective" and "Non-Perspective" Spatial Illusions', *British Journal of Psychology*, 69 (2) (1978), 233–41. Nor is this an isolated case: a casual dip into the relevant literature reveals the existence of variable susceptibilities, whether cross-cultural or individual, to other illusions, including, for example, the Ponzo and Ebbinghaus illusions.

[32] Compare R 4117 (17: 423), where Kant suggests that in the cosmological argument a necessary being is posited in order to bring to a close the series of *why*-questions arising from inquiring about the reasons for the existence of alterable things, thereby making the existence of alterable things fully comprehensible. Kant must mean that it is the fact of the *existence* of this being rather than the being itself that is taken by proponents of the cosmological argument to be the answer to the final envisaged *why*-question. I am grateful to James Kreines for this last point.

1.6 Some recalcitrant texts

Up to now I have for expository reasons been suppressing certain texts that complicate Grier's otherwise helpful picture. Like Grier, I have emphasized the point that for Kant, although transcendental illusion unavoidably tempts us to err, it does not force our hand.[33] But although such a view seems to be well supported by the texts, Kant does not always use the term 'transcendental illusion' for illusion in contrast to error. Indeed, in a number of places he uses the terms 'illusion' [*Schein*] and 'transcendental illusion' for preventable, false *judgment* or *belief*. I end this chapter with a brief look at these recalcitrant texts.

Consider, first of all, some passages in which Kant uses the term 'transcendental illusion' for something that we are able to remove or prevent. One such passage occurs in the Antinomies. In the course of discussing his general solution to these apparent conflicts of reason, Kant says:

> If one regards the two propositions, "The world is infinite in magnitude," [and] "The world is finite in magnitude," as contradictory opposites, then one assumes that the world (the whole series of appearances) is a thing in itself.... But if I take away this presupposition, *or rather this transcendental illusion* [*transcendentalen Schein*], and deny that [the world] is a thing in itself, then the contradictory conflict of the two assertions is transformed into a merely dialectical conflict.... (A 504–5/ B 532–3, emphasis added)

The suggestion that transcendental illusion can be removed appears again in the *Real Progress* of 1793:

> [In] the concept of the unconditioned in the totality of all mutually subordinated conditions.... there is need to remove that illusion [*Schein*] which creates an antinomy of pure reason, by confusion of appearances with things in themselves ... (*Real Progress*, 20: 311)

Each of these passages uses the term 'illusion' for something removable, and each suggests—the first more explicitly than the second—that this removable something is a belief (or presupposition) in whose rejection the resolution of the antinomies consists, namely, the belief in Transcendental Realism.[34]

One supposes that Kant may have found it natural to describe Transcendental Realism as an 'illusion' because he would have regarded the belief in this doctrine

[33] See: A 644–5/B 672–3; A 821/B 849; *Prolegomena*, 4: 328, and 28: 583.

[34] It is not clear that the content of transcendental illusion should in general be equated with the content of the belief in Transcendental Realism, but it is plausible that the instance of it associated with the cosmological Ideas should be.

as an especially deep, yet overlooked, feature of pre-critical philosophy. And yet, such talk is, strictly speaking, inconsistent with his official line on transcendental illusion; for Transcendental Realism consists not in an unavoidable illusion but rather in an avoidable belief. Kant's better thought would seem to be that Transcendental Realists are taken in by transcendental illusion (proper) in rational cosmology insofar as they fall prey to the illusion that the world exists as an absolute totality (the illusion, that is, that Transcendental Realism is true). This better thought is expressed in the first *Critique* when Kant says that in the antinomies 'Transcendental illusion has portrayed a reality to [the transcendental realist] where none is present' (A 501/B 29–30).

A second group of recalcitrant texts suggests that Kant sometimes uses the term 'transcendental illusion' to refer to the *product* of dialectical inferences rather than to their cause. Such a usage coheres with (what I have argued to be) Kant's view that the illusion that *D* is true is the product of certain other sub-illusions. But the usage, I think, indicates something that goes beyond this idea. At the opening of the Antinomies he says:

> We have shown in the introduction to [the Dialectic] that *every* transcendental illusion [*transcendentale Schein*] of pure reason *rests on* [*beruhe auf*] dialectical inferences. (A 405/B 432, emphases added)

And again, in the Ideal he says:

> [In the cosmological argument] speculative reason seems to have summoned up all its dialectical art so as to bring about the greatest possible transcendental illusion [*transcendentalen Schein*]. (A 606/B 634)

The context of this last remark makes clear that here Kant is referring to the untrustworthy inferences involved in the cosmological argument. He is personifying reason as an ancient dialectician, fluent in the deceptive art of creating illusion—and by 'illusion' here he seems to mean the body of conclusions of these untrustworthy inferences.

Further passages suggest that Kant occasionally uses the word 'illusion' to refer either to a judgment or to the 'taking' of one thing for another—a mental act that is itself naturally construed as a judgment. Consider, for example, the following three remarks:

> It is correctly said that the senses do not err; yet not because they always judge correctly, but because they do not judge at all. Hence truth, as much as error, *and thus also illusion* [*Schein*] *as leading to the latter, are to be found only in judgments*, i.e., only in the relation of the object to our understanding.
>
> (A 293/B 350, emphasis added).

> One can place all **illusion** [*Schein*] in the taking of a **subjective** condition of thinking for the cognition of an **object**. (A 396)
>
> [N]othing is more natural and seductive than the illusion [*Schein*] of taking the unity in the synthesis of thoughts for a perceived unity in the subject of these thoughts. (A 402)

In these passages Kant is speaking not of the illusion (in the sense of a seeming) *underlying* the formation of false dogmatic beliefs, but rather of the illusion (in the sense of an error) that results from the commission of a fallacy. This error, under its most general description, is that of *taking* something that is in fact merely subjective for something objective. As we will see, this description fits perfectly the error under discussion in the Paralogisms chapter, where Kant scrutinizes certain inferences that he describes as containing an illusion *within* themselves (A 396).

What these last-considered passages reveal, I think, is simply that Kant is engaging in equivocation. He uses the word 'illusion' sometimes for a kind of pervasive, seductive, and systematic *seeming*, and sometimes for a judgment, belief, inference, or phenomenon that is 'false' in one of the following senses: mistaken, invalid, specious. An adjectival occurrence of 'illusion' corresponding to this last usage occurs in a remark from Kant's metaphysics lectures. 'We will be able', he says,

> to secure morality and religion against the specious objections of speculative reason [*Schein Einwurfe der speculativen Vernunft*].
>
> (*Mrongovius Metaphysics*, 29: 781)[35]

Some of the passages just considered also suggest that Kant occasionally uses the whole phrase 'transcendental illusion' equivocally. In particular, when the 'illusion' (in the sense of 'error') has its source in 'transcendental illusion' (a seeming), Kant is prepared to term it 'transcendental illusion'.

These complications of the initial picture help to explain why it should have taken so long for Grier's point to be made. Her insight arose because she noticed a central strand of Kant's thought that tracked a certain sub-set of occurrences of the terms 'illusion' and 'transcendental illusion'. By identifying this strand and emphasizing its importance, Grier has revealed the coherence of Kant's position on transcendental illusion and opened up the Dialectic as never before.

As we noted at the beginning of this chapter, Kant characterizes 'transcendental dialectic' as the 'critique of illusion'.[36] We have now arrived a clearer view of what this critique involves, namely, at minimum, the detection of transcendental

[35] Kant's note-taker capitalizes '*Schein*' when it occurs as an attributive adjective. Compare: '*Schein Weisheit*' (29: 766).

[36] R 5063, 18: 76–7.

illusion—understood as a pervasive and permanent illusion of reason that systematically tempts the unwary philosopher into dogmatic error (A 297/B 354). I would venture that, beyond this, the critique of illusion also involves some account of the *workings* of this illusion: some account along the lines, perhaps, of the transitions leading from [1] to [4] in §1.2. Finally, it involves the tracing of reason's itinerary as it works through its natural dialectic. We begin our examination of that itinerary with a consideration of the claims and prospects of rational psychology.