3 The Antisceptical Argument of the Fourth Paralogism

In the preceding chapters we saw that the *Critique* in general and the Fourth Paralogism in particular, if read with the appropriate hermeneutical tools, are committed to a kind of idealism that despite appearances can and should be sharply distinguished from phenomenalism. We saw that Kant's idealism has the resources to rebut standard criticisms regarding its nature and its foundations that aim at threatening its plausibility. In this sense, transcendental idealism can support the weight of being the fundamental premise of the antisceptical argument of the Fourth Paralogism. The present chapter spells out this argument. The key lies in reinterpreting the sceptic's challenge from the perspective of transcendental idealism (properly understood). From this standpoint, the sceptic's hypothesis – the world could be a consistent systematic hallucination – becomes a metaphysical question that can be ruled out as illegitimate. Specifically, it can be treated as a vain inquiry into the nature of the thing in itself.

On my reading, the Fourth Paralogism establishes this result through the following steps. To begin with, by placing the critique of Descartes in the Paralogism chapter, Kant is implicitly suggesting that there is a connection between Descartes's scepticism and transcendental illusion, the force that motivates the drive towards the unconditioned that is critically analysed in the Transcendental Dialectic. In an attempt to determine this connection – which Kant never spells out – I will argue that the thesis of the epistemic superiority of inner over outer knowledge – the crucial premise of Descartes's scepticism – rests largely on the mistake that infects rational psychology, namely, it misinterprets a subjective condition of thought as a characteristic of an object (the subject, understood as a particular determinate entity).[1] Second, the same

thesis of the superiority of inner over outer knowledge is further weakened through Kant's proof that not only the perception of our mental states and of our existence, but also the perception of external objects, is immediate. Thirdly, Kant deals with the crucial problem of hallucinations or delusions caused by the imagination. On the one hand, he introduces a criterion for detecting these forms of delusions and thereby distinguishing them from genuine experience. This criterion turns on the idea that a particular hallucination can be detected by its failure to harmonize with the laws of experience in general. Yet on the other hand, he dismisses the possibility that the entire world could be a kind of systematic or lawful hallucination (a kind of hallucination that obviously would not be detected by using the criterion) with the argument that the very functioning of the imagination (the faculty that would produce this systematic hallucination) presupposes the existence of external objects as the sources of the material on which the imagination exercises itself.

For good reasons, this last step has been rejected by virtually all interpreters. So I will propose replacing it with a reflection on the very meaning of this systematic hallucination hypothesis – which I take as identical with the Evil Genius hypothesis – within the framework of transcendental idealism. Using materials that Kant himself suggests in the Fourth Paralogism and elsewhere, I will show that wondering whether the cause of all our representations of external objects is itself external or merely internal amounts to wondering about the nature of the thing in itself. This wonder – which is the essence of the sceptic's hypothesis – is illegitimate in the context of transcendental idealism; furthermore, the possibility that this thing is some spiritual entity such as the Evil Genius can even be granted by Kant because it is not the existence of this 'thing' that a transcendental idealist wants to secure. Whatever that 'thing' is that affects us, be it a spiritual or a material entity, we represent it as a world in space through the form of outer sense. The existence of this empirical world is all that concerns the transcendental idealist.

Thus my answer to scepticism will not rest on detecting contradictions in the sceptic's position. Nor will I attempt to identify a premise to which the sceptic is committed (such as the reality of inner experience) in order to show that it presupposes something that the sceptic denies or doubts (such as the reality of outer experience). In the proof I offer, the 'game of idealism' will *not* be 'turned against itself' as Kant thought he was able to do with his 1787 Refutation of Idealism. I will not even

be arguing, as Carnap did, that empirical idealism is meaningless because it falls outside what can be verified. My opinion – which I believe also would have been Kant's – is that empirical idealism is a philosophical stance with a legitimate and clear meaning. My strategy, then, will be to uncover the metaphysical nature of the sceptical hypothesis, which is obscured by its arising from a seemingly innocent generalization of normal cases of hallucinations. In particular, my argument will revolve around reducing scepticism to an inquiry into the nature of the thing in itself.

The steps to my argument dictate the structure of this chapter. In the first section, I explain why Kant views Descartes's scepticism as a particular case of transcendental illusion. In the second section, I focus on the thesis of the immediacy of outer perception, which is the first ingredient in Kant's overall refutation. Since this thesis depends largely on the arguments of the Metaphysical Exposition in the Transcendental Aesthetic, I will presuppose the result of my analysis in the preceding chapter. So that readers will be able to appreciate the strength and originality of Kant's position, I will be proposing a comparison with Reid's affirmation of the same thesis (our outer experience is not mediated by 'ideas'). In the third section, having introduced Kant's criterion for distinguishing hallucinations from real experience, I present the antisceptical argument; then in the fourth section I rebut what I consider to be a crucial but also last-resort counter-objection open to the sceptic. The chapter closes with some remarks on Kant's supposed abandonment of the Fourth Paralogism argument in the Second Edition.

The Fourth Paralogism and Transcendental Illusion

The first question the interpreter confronts in dealing with the Fourth Paralogism is, quite simply, its location. Why does Kant place the critique of scepticism among the Paralogisms? Given the architectonic of the *Critique*, the Fourth Paralogism corresponds to the modal categories and these 'have the peculiarity that, in determining an object they do not in the least enlarge the concept to which they are attached as predicates. They only express the relation of the concept to the faculty of knowledge'(A219/B266). This means that they assert whether an object – whose concept they do nothing to determine – is existent, or necessary, or merely possible. Since the Cartesian sceptic does not doubt the appropriateness of our concepts of things, but rather whether (external) things exist, the placement of the critique of this kind of scepticism in

the Paralogisms seems to be rather natural. Moreover, rational psychology – the science that is under attack in this section of the *Critique* – is in the main a Cartesian project, and this by itself would justify such placement. Besides these two reasons, however, I think there is a more profound one, which has to do with the connection between transcendental illusion and a basic premise of scepticism – namely, the superiority of inner over outer knowledge.

At the end of the 1781 Paralogisms chapter, Kant offers us one of the general definitions of transcendental illusion contained in the Dialectic.[2] He contends that 'all *illusion* may be said to consist in treating the *subjective* condition of thinking as being knowledge of the *object*' (A396). In the Paralogisms this means taking the logical predicates contained in the notion of a thinking subject as features of a particular, determined object (the subject understood as a determinate entity). Kant states this very clearly in the Second Edition: 'The logical exposition of thought in general has been mistaken for a metaphysical exposition of the object' (B409). In the first three Paralogisms the logical features of the thinking subject in general that are mistaken for objective features are substantiality, simplicity, and personality. But what, precisely, is the logical feature that rational psychology mistakes for an objective feature in the Fourth Paralogism?

The Fourth Paralogism corresponds to the modal group of categories – in particular, to the category of existence.[3] So one would expect that the feature in question is the very existence of the subject. Since a subjective condition of thinking is, obviously, that the thinking subject exists, the logical feature mistaken for an objective feature should be the very existence of the subject. At first glance, however, the existence of the subject is not the feature that Kant seems to have in mind. In the introductory 'topic' of rational psychology, where, in correspondence with the four groups of categories, he lists the logical features of the thinking subject that rational psychology mistakes for objective features, he characterizes the feature corresponding to the modal group of categories as follows: 'It [the soul] is in relation to *possible* objects in space' (A344/B402). As the reader discovers in the Fourth Paralogism, the objects with which the soul is in relation are merely *possible* because their existence, ascertained through an inference from effect to cause, is doubtful.

Although Kant attempts to present this attribute as a feature of the soul (how it relates itself to objects in space), it prima facie seems to be a feature that concerns more the objects and the way in which they exist

rather than the soul itself. However, this impression would be mistaken. The existence of objects is merely possible because of the particular way in which the soul ascertains their existence; just as importantly Kant contrasts this way with the way in which the soul ascertains its own existence. Indeed, the absolutely certain access that the soul has to its own existence seems to be the background against which he contrasts and evaluates the merely possible existence of external objects. Thus to say that the soul is in relation to objects in space that are merely *possible* because their existence is inferred is an indirect way of emphasizing the immediacy and certainty through which the existence of the soul can be ascertained.

This is confirmed, in the final section of the First Edition version of the Paralogism chapter, by the approach Kant takes to characterizing the feature of the soul corresponding to the Fourth Paralogism. He states that the thinking 'I' regards itself 'as the correlate of all existence, from which all other existence must be inferred' (A402), and adds that the thinking subject knows '*the unconditioned unity of existence in space, i.e.* that it is not the consciousness of many things outside it, but the consciousness *of the existence of itself* only, and of other things merely as its *representations*' (A404). These expressions are quite obscure, yet Kant's point is clear enough: what the rational psychologist assumes he is able to determine through merely logical means – that is, on the basis of the mere 'I think' – is '*the existence of itself* [the subject] only.' Access to the existence of other beings occupies an inferior epistemic level because those beings' existence is inferred from the representations that lie within the subject. The existence of the subject, or 'unconditioned unity of existence,' is thus the real feature at stake in the Fourth Paralogism. This feature serves as the reference point – as the paradigm of the unity enjoyed by possible objects in space. This explains why Kant characterizes it as the unconditioned unity of existence '*in space.*'

Thus the original error analysed in the Fourth Paralogism is not that the soul regards outer objects as merely possible. This is a mere consequence. The original error lies in taking one's existence in that paradigmatic manner, which, in turn, is a consequence of assuming that one's existence can be ascertained through merely logical means. In the Fourth Paralogism, transcendental illusion is precisely the force driving the inference from this particular condition of thinking (existence of the subject) to the existence of the subject as an objective fact. Since this existence is analytically inferred from the 'I think,' it is absolutely certain and, as such, is in contrast to the doubtful (or possible) existence of

external objects. The analytic nature of the inference that leads to the existence of the subject engenders an enormous epistemic gap between the affirmation of the existence of the subject and that of external objects. Indeed, if we concede that our existence as subjects can be ascertained in this analytical way, knowledge of our existence will differ from knowledge of the existence of outer objects in *kind*, not just in degree of certainty. The former enjoys a priori status; the latter can only be a synthetic a posteriori claim.

By placing the refutation of scepticism in the Fourth Paralogism chapter, Kant is suggesting that the first step in refuting the sceptic entails showing that this difference in kind is merely apparent. More precisely, by revealing how the idea of an analytical inference from the logical description of the subject of thought to the existence of the subject is merely illusory, Kant is removing one of the grounds on which the superiority of inner over outer knowledge rests. Once the mistake that lies at the heart of this inference is detected and it turns out that even the cognition of my existence presupposes an empirical intuition (perception), inner and outer knowledge are no longer considered different in kind.

Note that Kant, on my reading, does not establish this point through his actual discussion of the Fourth Paralogism, but merely through the placement of his criticism within his general treatment of rational psychology. This placement suggests – albeit implicitly – that the same ipostatization of the features of a merely logical subject (the 'I' of apperception) into a noumenal subject that Kant detected in the first three Paralogisms is also key to the criticism of the fourth. Once the idea of an intellectual access to the existence of a subject is replaced with the idea of an empirical access to ourselves, Kant can introduce his crucial idea – which he explicitly presents in his discussion of the Fourth Paralogism – that both the existence of myself and that of an external object are 'proved in the same manner' (A370) and that 'the only difference is that the representation of myself, as the thinking subject, belongs to inner sense only, while the representations which mark extended beings belong also to outer sense.' (A371) In fact, the critique of the Fourth Paralogism *assumes* that our existence must be *perceived* in inner sense. Indeed, the first premise of the Fourth Paralogism is that 'my own existence is the sole object of a mere perception' (A367), or, equivalently, that Descartes was justified in limiting 'all perception' to the proposition 'I, as a thinking being, exist'.' (A367). Since it is hardly trivial that Descartes took the cogito as resting on a perception, Kant is here

presenting the Cartesian position as already reinterpreted and criticized. And the criticism is precisely that there is no access to my existence without empirical intuition – specifically, the perception of my thoughts in inner sense.[4]

If this analysis explains why Kant placed his critique of scepticism in the Fourth Paralogism and what the connection is between scepticism and transcendental illusion, it opens up a new twofold problem. On the one hand, the very idea that we become aware of our existence necessarily through a perception is ambiguous. In the *Critique* Kant presents two different modes of self-consciousness – one through inner sense, the other through apperception. Each is based on a 'perception,' but in a very different way. The problem for us is simply to ascertain which mode Kant is assuming in the Fourth Paralogism.

For Kant, through apperception we can become aware of our spontaneity – that is, our activity as thinkers. But this activity cannot be captured by an intuition. Indeed, any intuition would give us only a *thought* and not the very activity of *thinking*. As Kant puts it in a footnote in the Transcendental Deduction: 'Since I do not have another self-intuition which gives me the determining in me (I am conscious only of the spontaneity of it) prior to the act of *determination*, as time does in case of the determinable, I cannot determine my existence as that of a self-active being; all I can do is to represent to myself the spontaneity of my thought, that is, of the determination' (B158n).

Moreover, according to Kant, in this non-sensible consciousness of the activity of thinking one can also find an awareness of existence: 'In the synthetic original unity of apperception, I am conscious of myself, not as I appear to myself, but only that I am. This *representation* is a *thought*, not an *intuition*'[5] (B157). The content of this 'pure' self-consciousness cannot be provided by an (empirical) intuition, yet perception is still necessary for it. As Kant clarifies in the Second Edition version of the Paralogism, I need an empirical material (what perception provides) for the very exercise of the faculty of thinking. But this material functions merely as the *occasion* of the act of thinking. In fact, the object of the perception that is necessary for this act need not even be the subject itself or one of its inner states. It could be anything, because independently of its object, any perception can play the role of occasioning the activity of thinking. Once this activity is occasioned, although we can really *know* only the products of this activity, we can be *aware* of the activity itself that generated those products. In turn, once we become conscious of this activity, we can also become aware of the

existence of an agent that does the activity. In this sense, we are aware of the existence of the subject as thinking agent.

Self-knowledge through inner sense is based on perception in a manner very different from self-consciousness through apperception. In the former case, perception is not a mere occasion, as we perceive the mental states that are successively given in inner sense. It follows that perception is what provides the very content of our cognition. Since we will be dealing with the details of this kind of empirical self-knowledge in the following chapter, we can limit ourselves here to clarifying the role that perception and inner sense play in this kind of cognition.

This in itself is sufficient to solve our problem of determining the mode of self-knowledge with which Kant is operating in the Fourth Paralogism. Kant's explicit reference to inner sense tells us that he is operating in the Fourth Paralogism with the first, 'thick' mode of self-knowledge – that is, inner experience. So, for example, he claims: 'External things as well as myself exist upon the immediate witness of my self-consciousness. The only difference is that the representation of myself, as the thinking subject *belongs to inner sense* only, while the representations which mark extended beings belong also to other sense' (A371, my emphasis).

But if it is clear which notion of self-consciousness Kant is operating with, it is less clear whether his main point holds true – that both the existence of myself and that of external objects are 'proved in the same manner,' (A370) – considering that he himself acknowledges the other, non-sensible mode of becoming aware of our existence. In fact, the idea that in apperception the existence of the subject is already given (and given in a non-sensible manner) seems to contradict the presupposition that we thought Kant wanted to rule out by placing his critique of scepticism among the Paralogisms. The idea of a non-sensible access to our existence seems to make self-consciousness again different *in kind* (rather than in degree of certainty) than the consciousness of outer objects.[6]

One way to solve the problem is to understand that the 'I' whose existence is contained in apperception is not what Descartes (and the rational psychologist of the Fourth Paralogism) thought he was able to infer from the consciousness of thinking. The 'I' that arises from the radical doubt is a *res cogitans*. As we have seen, however, through apperception we become conscious of an activity, not of a thing. The thought of this activity obviously includes also the thought of an agent. But by no means can we take this as cognition of a determined

entity or (more simply) of a thing. It is nothing but the pure thought of the subject of this activity. Through this thought, the number of the 'items' of which the world is composed is by no means increased. In order to come to know the existence of myself as one of the 'things' that constitute the world, I require a perception, no longer as mere occasion but as the *vehicle* that provides the determinations of that particular 'thing' (the empirical 'I'). It follows that for Kant, notwithstanding his misleading expressions, the only way to ascertain the existence of ourselves as actual 'things' is through inner experience. And this tells us that the result established by positioning the refutation in the Paralogism – that is, no intellectual access to our existence, and therefore no difference in kind between inner and outer experience – is not undermined by Kant's doctrine of apperception as a mode of self-consciousness.

A last point needs to be clarified. By showing that we cannot analytically ascertain our existence, as the rational psychologist contends, Kant is not saying that access to our existence is dubious. The fact that we know the existence of ourselves as 'objects' through inner sense does not mean that my existence becomes dubious because it is open to the delusion of the imagination. In criticizing the cogito, Kant never meant to refute Descartes's point that inner experience is immune to the kinds of delusion to which outer experience is vulnerable. The point is simply that this undeniable advantage of inner knowledge need not be taken as the result of a presumed 'intellectual' access to our own existence. Indeed, what the actual discussion of the Fourth Paralogism (as opposed to its placement) is meant to convey is that the *immediate* relation between the mind and the object that is rightly predicated in the case of inner knowledge should *also* be predicated on outer knowledge, notwithstanding the undeniable delusion that it suffers. Briefly put, the point is not to lower inner knowledge to the level of outer knowledge, but rather to elevate (as much as possible) outer knowledge to the level of inner knowledge. This result can never be achieved if we first do not free ourselves of the mistake (motivated by transcendental illusion) that inner knowledge enjoys a kind of *'intellectual* immediacy' denied to outer knowledge. Once it is shown that the access we have to our own existence, albeit certainly immediate, is still of a sensible kind, a preliminary obstacle on the way to re-establishing a certain parity between inner and outer knowledge is removed. The next antisceptical step is precisely to show that outer knowledge enjoys the same kind of (sensible) immediacy as does inner knowledge.

The Immediacy of Outer Perception

Before asking how Kant manages to ground the immediacy of outer perception, we must confront a preliminary problem, one that hinges on the fact that the very notion of an 'immediate perception' (be it inner or outer) is problematic within Kant's system. For Kant, only intuitions are immediate whereas concepts and judgments entertain a mediate relation with the object. Kant makes this point quite clearly in the section of the Analytic devoted to the logical employment of the understanding: 'Since no representation, save when it is an intuition, is in immediate relation to an object, no concept is ever related to an object immediately, but to some other representations of it, be that other representation an intuition, or itself a concept. Judgment is therefore the mediate knowledge of an object, that is, the representation of a representation of it' (A68/B93).

The problem with the notion of immediate perception is that in many places in which Kant defines perception, he takes it as involving a kind of judgment. In fact, what distinguishes a perception from a mere sensation is that only the former contains a conceptual synthesis. For example, in the B version of the Metaphysical Deduction, Kant writes that 'all synthesis, therefore, even that which renders perception possible, is subject to the categories' (B161). But perception does not involve only pure concepts. In general, perception is presented as a sensation to which a concept is applied, be it pure or empirical. An objective perception, as opposed to a merely subjective one (that is, to a sensation), is even identified with knowledge.[7] Since knowledge obviously requires the application of empirical concepts as well as pure ones, and since this application is nothing but an act of the faculty of judgment, it turns out that a perception is a judgment in which some (pure or empirical) concepts are applied to sensible data. But if a perception is a kind of judgment, and if a judgment is a 'mediate knowledge of an object,' how can Kant consistently talk of an *immediate* perception?[8]

The solution to this difficulty turns on the precise determination of Kant's notion of immediacy. In the context of the Fourth Paralogism, immediacy means that a temporal or spatio-temporal *indeterminate* object is given to the mind without the medium of a third thing. The understanding (taken here as what synthesizes the sensible material through pure and empirical concepts) is presented with an object that is already given to the mind as either temporal or spatio-temporal. The presentation of this indeterminate temporal or spatio-temporal object,

which is the autonomous contribution of sensibility, is what informs me about the existence of 'something' in space or in time. As Kant puts it in the body of the Fourth Paralogism: 'It is sensation ... that indicates a reality in space or in time, according as it is related to the one or to the other mode of sensible intuition'(A373–4). Judgment – determinative judgment, to be precise – enters the picture when this already either temporal or spatio-temporal indeterminate 'something' has to be brought under a concept. In this way, what used to be merely a temporal or spatio-temporal indeterminate 'something' becomes a determinate object. By subsuming this object under a concept, judgment relates a representation (the concept) not to the object itself, but to the object that is already organized according to the forms of sensibility. For this reason it is in a mediate relation to the object. But the existence of an indeterminate object was already given through the apprehension of this temporal or spatio-temporal material, and it was given immediately. Thus the mediacy through which the material is organized under pure and empirical concepts does not contradict the immediacy through which the not yet determined object is given to sensibility. It follows that Kant can consistently talk of an immediate perception even if the latter involves judgment.

Moving to the *foundation* of the immediacy of outer perception, the question becomes this: How is the existence of an (indeterminate) object given to my sensibility *immediately*? The minor premise of the Fourth Paralogism reads: 'All outer appearances are of such a nature that their existence is not immediately perceived, and that we can only infer them as a cause of a given perception' (A367). Much of the anti-sceptical argument of the Fourth Paralogism turns on whether Kant succeeds in refuting this claim.

Kant uses for this refutation two basic tools. One is the result established in the Transcendental Aesthetic, in which space is reduced to the form of our outer sense – a form that presents itself to the mind as a pure intuition. Implicit in all this is the standard definition of intuition as what 'refers immediately to its objects' (A320/B377). The other is the criticism of the transcendental realist's model of perception.

Starting with the last point, Kant contends that the idea cherished by the sceptic – that the existence of outer things is *not* immediately perceived – is a direct and perverse outcome of transcendental realism.[9] As he puts it: 'Transcendental realism ... inevitably falls into difficulties, and finds itself obliged to give way to empirical idealism, in that it regards the objects of outer sense as something distinct from the senses

themselves, treating mere appearances as self-subsistent beings, existing outside us. On such a view as this, however clearly we may be conscious of our representation of these things, it is still far from certain that, if the representation exists, there exist also the object corresponding to it' (A371).

Transcendental realism assumes that the objects of our experience are things in themselves. As such, they are what they are independently of our senses. There is nothing in the objects of our experience that they have by virtue of our intuition of them. Given this model of the mind/object relation, Kant argues that scepticism is unavoidable.[10] What is presented immediately to the mind is *never* the object, but always a copy or representation of it. In sharp contrast, for a transcendental idealist space is not an objective property that must be 'picked up' (somehow) by the mind; rather, it is a form through which we first become aware of objects other than ourselves. The problem of scepticism arises precisely when some sort of 'picking up' is introduced as the model of a cognitive act. Kant's idealism can be construed precisely as a denial of this idea.

Once again, everything turns on the status attributed to space in the Transcendental Aesthetic. Recall, there it was shown that the way in which objects are what they are is *not* independent of our sensibility. For the transcendental idealist, space is not an objective property, but the form that makes possible the very representation of something other than myself. As such, space infallibly enters the constitution of an object of experience other than the knowing subject. The immediacy through which objects other than myself are given is thus guaranteed by the fact that space, rather than being a feature among others that must be derived from experience, is in fact the subject's insufficient yet necessary contribution towards the constitution of an external object. To use the (too) common metaphor, on Kant's account the immediacy of outer perception is guaranteed by the fact that the mind 'shapes' whatever affects my outer sense.[11] To perceive an object means to constitute it as spatial, and obviously there is no room for mediacy here.

Perhaps this crucial idea of the immediacy of outer perception can be clarified further by concentrating no longer on the status of space as *form of intuition*, but rather on the status of the representation of space as *formal intuition*. Kant argues that 'space and time are represented *a priori* not merely as forms of sensible intuition, but as themselves *intuitions* which contain a manifold [of their own]' (B160). Moreover, this manifold has priority over the representations of particular objects con-

tained in it.[12] As Kant puts it: 'We can never represent to ourselves the absence of space, though we can quite well think it as empty of objects' (A24/B38–9). Even if there were no objects, we would still intuit an external framework. Although we cannot intuit the object 'space' as an infinite container devoid of determinate objects, we can intuit an external horizon – let us call it an externality – that is waiting to be filled with determinate objects. There can be no question that this external framework is given to us immediately. In fact, the Transcendental Aesthetic shows that it is given to us as an a priori intuition, and any intuition – let alone an a priori intuition – is by definition the cognitive act that allows the mind to make direct contact with the object.[13]

Obviously, all of this merely shows that the pure, formal intuition of space is given immediately. What can we say about the spatial *objects*, with all their empirical determinations? If space, as a pure horizon, is given immediately, and if external objects are determinations of this external horizon, they are themselves immediately given. In occupying parts of this horizon, empirical objects borrow the immediacy through which the horizon itself is given. How could something that merely delimits a horizon immediately given be itself not given in the very same way? Hence there is no inference in our perceiving an external object. Its presence is immediately grasped because space, being an intuition with its own (pure) manifold, is immediately given in our cognition, and the object itself is simply a determination of this given.[14]

A quite paradoxical, but I think helpful, way of expressing the same point is to say that on Kant's account, the mind 'reaches' or 'touches' the external object. The reason is again that external objects occupy a place on a horizon that is pre-given to the mind. It is as if the subject, thanks to the pure intuition of space, could consider each spatial object as something that, as determination of its horizon, acquires a 'closeness' to itself. This obviously does not mean that the object ends up being 'in' the mind as a mental entity. The idea is quite the opposite. The pure intuition of space allows the subject, as it were, to go outside of itself and be in direct contact with everything that occupies a determined place in its spatial horizon.[15]

Reid's Realism

There is, of course, a lot of contention over Kant's certainty that a transcendental realist is committed to an indirect model of perception. Kant

seems to operate with the model of classical empiricism, yet he offers no explanation why, for example, direct realism is not even a possibility. One would think that there is nothing incoherent in the position according to which (1) we know things as they are in themselves, and (2) we have direct access to them – that is, access not mediated by a third entity such as a mental entity of some sort. So it seems that a transcendental realist is by no means committed to the gap between mind and object that Kant confidently assumes.

In order to fully appreciate the strength and originality of Kant's defence of the immediacy thesis, and to ground his assessment of transcendental realism as necessarily committed to the sceptical threat, it would be instructive to discuss briefly Reid's theory of perception. Through his attack on the theory of ideas, Reid attempts to reach roughly the same result as Kant: the relation between the mind and external objects is *not* mediated by a mental entity. Interestingly, Reid's argument rests on an intuition very similar to the one exploited by Kant: he considers space a necessary component of our outer perception *without being derived from outer perception itself*. In this way he champions the direct realism that we introduced earlier as a possible counter-example to Kant's general characterization of the transcendental realist.

In his critique of Berkeley and Hume, Reid sets up an *experimentum crucis* to test the soundness of scepticism: either space is an 'idea of sensation,' by which he means something that can be derived from the experience of objects, or it is not. If it is, then skepticism is unavoidable because every experience of a spatial object is mediated by an idea that is supposed to resemble the spatiality of things. If it is not, space turns out to be an irreducible brute datum of our experience that is immediately given to us *and that cannot be put at risk out of the supposed necessity that we bring it to our consciousness through a representation*. Any attempt to show that this idea is derived by the senses (Reid deals specifically with Locke's attempt to derive the idea of space from the sense of touch) is actually circular because the 'externality' of these experiences is always presupposed.[16] It follows that space and the things in it are not given through an idea in the mind, but immediately.[17]

There are similarities between Kant and Reid regarding the non-empirical origin of the representation of space, but there is also a crucial difference: Reid does not explain the surprising fact that space is not derived from experience and yet is a necessary component of each outer perception. He seems to lack the resources to explain how an

affection, which supposedly only brings what is in the affecting thing – can make us conceive immediately something that, *ex hypothesi*, is not in the thing itself, namely, its spatial location. He admits as much in a revealing passage: 'How a sensation should instantly make us conceive and believe the existence of an external thing altogether unlike to it, I do not pretend to know.' All Reid is willing to say is that, by a quite mysterious law of nature, 'such a conception or belief [space] constantly and immediately follows the sensation.'

From a Kantian perspective, Reid's position fails to recognize that space is a 'form of sensibility' and that it functions as a condition for the representation of objects other than ourselves. Lacking this insight, Reid is open to the charge that even if we admit his point that no sensation can originate the representation of space, two options are still open to the sceptic. One is to say that the idea of space is an innate idea, which is obviously compatible with the point that all we actually perceive are ideas (innate and non-innate). The other is simply to blame Reid for asking us to accept this mysterious mechanism according to which the idea of space arises from an affection, whereas the object that affects our sensibility is not supposed to provide the feature of spatiality. Why should the sceptic accept the immediacy of outer experience if it ultimately rests on a mechanism that Reid himself considers 'utterly inexplicable'?[18]

These two objections are inevitable consequences for the peculiar kind of realism Reid champions. If we lack the fundamental idea that objects are organized by the forms of our sensibility, but at the same time want to deny the empirical origin of some objective properties, we can only conclude either that we have an innate idea of space or that we are confronted with an inexplicable mystery. An idealist, conversely, suffers neither of these consequences. To begin with, the idea that space is a form of our sensibility has very little to do with the notion that it is an innate idea. As we saw in our critique of Guyer in the preceding chapter, Kant's notion of a condition of the possibility of experience is very different from that of an innate idea (as well as less vulnerable to the skeptical threat). Moreover, the idea of space as a form of sensibility at least explains the peculiar status of the representation of space. All of this suggests that Kant's idealism manages to ground the immediacy thesis in a much more convincing way than any realist approach, regardless of its good intentions.

Is the significance of this result really limited to the peculiarities of Reid's version of realism, or we can rather generalize our findings to all

forms of realism, including direct realism? In order to avoid the difficulties Reid confronts, the direct realist could hold that space *is* a property derived from the external objects, but insist that the mind perceive the spatiality of things without any mediation. On this reading, the mind would be presented directly with space, which would therefore be considered along with all other properties of external objects. As I have said, there is nothing inherently incoherent in this position, but the realist needs to show that the arguments of the Transcendental Aesthetic are flawed. There Kant proved that the representation of space is not an 'empirical concept' (A23/B38), mainly because any external representation presupposes it. Thus even if we were to concede to the direct realist that the whole of our perception functions as he suggests – that is, through some sort of unmediated access to the property of things – the property 'space' would nonetheless constitute an important exception, as long as the first two arguments of the Metaphysical Exposition hold. Thus, either the direct realist convinces us that these arguments are incorrect, or Kant's characterization remains valid regarding the close relation – as inevitable as undesired – between realism (of all kinds) and scepticism. Needless to say, this shifts the burden of the proof onto the Transcendental Aesthetic, and we can here simply refer to our defence in the preceding chapter.

We now move on from the foundation of the immediacy thesis to its significance. The obvious consequence of the immediacy thesis is that the superiority of inner experience over outer experience is weakened (although, as we shall see, not completely erased). The intuition of my inner states (and thus the consciousness of my existence) and the intuition of outer objects are both immediate: 'External things exist as well as I myself, and both, indeed, upon the immediate witness of my self-consciousness' (A371). As Kant puts it: 'In order to arrive at the reality of outer objects I have just as little need to resort to inference as I have in regard to the reality of the object of my inner sense, that is, in regard to the reality of my thoughts. For in both cases alike the objects are nothing but representations, the immediate perception (consciousness) of which is at the same time a sufficient proof of their reality' (A371).

The sheer superiority of inner sense, to which Kant was still ready to subscribe in the late 1770s, is at last denied. The parity on which he insists, however, is open to an obvious objection. Even if we accept the idea that when I see a genuine external object, this object is perceived immediately, it is still the case that sometimes I merely *seem* to see such an object, but I am actually hallucinating. It is a fact about human nature

that we sometimes see things that on closer scrutiny do not exist. But these epistemic failures seem to affect only outer knowledge and leave inner knowledge completely untouched. Here Descartes's famous considerations in the Second Meditation hold. If I am conscious of a certain mental state (seeing the table), then, by that very fact, that mental state exists. Briefly put, the existence of a mental state coincides with its being represented. More importantly, *I* must exist in order to have that representation. Whereas in outer knowledge I may be wrong about the very existence of the object of my perception, the mere appearance of a representation in my mind assures me of the existence of myself as the being that entertains or thinks that representation. It follows that despite the immediacy of outer perception, inner knowledge still enjoys a certain degree of epistemic superiority over outer knowledge.

It is important to realize that Kant never meant to deny *this* superiority. When he contends that matter is 'proved in the same manner as the existence of myself as a thinking being is proved,' he only means to deny that there is a special kind of *intellectual* access to my existence (the mistake at the basis of the Fourth Paralogism of rational psychology). But this is perfectly compatible with the idea that our *sensible* access to ourselves and to our states is infallible. In other words, the cogito, if properly understood as based on the perception of my mental states leading to the knowledge of the empirical 'I,' is perfectly correct for Kant. The idea of denying the existence of this empirical 'I' is indeed for Kant simply 'an absurdity' (B71). Even merely doubting such existence is out of the question. As he points out in the Fourth Paralogism: 'There can be no question that I am conscious of my representations; these representations and I myself, who have these representations, therefore exist' (A370).

But if Kant holds both that the access to my representations and to myself is infallible and that the access to outer things is open to 'delusion of the imagination' such as hallucinations, it seems that the very notion of immediacy of outer perception is again put at risk. For one is tempted to infer from the existence of hallucinations that the normal relation between the mind and an outer object is mediated by a mental representation. The only difference between hallucinations and normal cases of perception would be that only in the latter case does an object actually correspond to the representation. This is indeed the reasoning on which scepticism is based. But this reasoning is clearly based on a non sequitur. The existence of hallucinations says nothing about how I perceive genuine external objects. Indeed, all that the immediacy thesis establishes is

that *if I ever perceive a genuine external object*, I perceive it immediately. This is clearly compatible not only with the existence of sporadic hallucinations, but also with the possibility that the entire world is a systematic hallucination. In that case, our ability to perceive genuine external objects immediately would simply never be actualized.

The result of this analysis is that the immediacy thesis *by itself* is not sufficient to refute the Cartesian sceptic. In fact, the immediacy thesis and the Evil Genius hypothesis seem compatible. Nonetheless, the immediacy thesis establishes the apparently minimal but (as we shall see) ultimately crucial point that if we were ever to perceive a genuine external object, it would be immediately. Indeed, to make this point is already sufficient to refute the sceptic's idea that even in cases of genuine experience, we perceive them mediately (through a mental entity). In the following sections we shall see that only if this idea is rejected can the refutation of the sceptic get off the ground. Though modest, the immediacy thesis plays precisely this dialectical role in the overall refutation.

The Refutation of the Sceptic: The Official Strategy

Kant's standard attempt to move beyond the immediacy thesis is found in a passage in which he acknowledges clearly the possibility of illusory representations to which no objects correspond, as in hallucinations or dreams: 'From perceptions knowledge of objects can be generated, either by mere play of imagination or by way of experience; and in the process there may no doubt arise illusory representations to which the objects do not correspond, the deception being attributable sometimes to a delusion of imagination (in dreams) and sometimes to an error of judgment (in so-called sense-deception)' (A376).

Now if notwithstanding the immediacy thesis, delusions of imagination are possible, it seems that the sceptic can easily raise his usual question: how do you know that what you have just acknowledged can happen occasionally does not happen systematically? To be sure, Kant provides a criterion for distinguishing such delusions from experience. But as he himself seems to realize, this criterion will turn out to be at best superfluous.

The criterion for recognizing delusions is the following: 'To avoid such deceptive illusion, we have to proceed according to the rule: '*Whatever is connected with a perception according to empirical laws, is actual.*'[19] Perhaps an example will help elucidate Kant's point. Let us assume that, being particularly afraid of the dark and finding ourselves in a dark

room, we are so overwhelmed by fear that we believe we see a threatening individual. However, we are lucky enough to find a light switch while we are still 'seeing' this individual. When the light goes on, we find that our room contains no dangerous company. We are led to consider our past 'perception' as a hallucination because to take it otherwise would commit us to giving up the empirical law (actually an application of it to this particular situation) that declares that bodies cannot 'go out of existence' or 'disappear' from our sight with such rapidity. In other words, we take our vision as a hallucination because it does not cohere with the rest of our well-tested and usually reliable empirical laws.

Interestingly, Kant considers the deception in question and the 'provision against it' – the criterion we have just seen – to affect both dualism and idealism. By these two terms he means respectively his position and the position of his opponent.[20] One might wonder why Kant argues that idealism should be affected by this provision. Why should the sceptical idealist be interested in distinguishing between real experience and hallucinations, if part of his point is that there is no certain way to distinguish between the two?

The answer is that even the sceptic must account for the regularity and coherence of experience. In other words, even if our entire experience is 'really' just a trick of the imagination or of some Evil Genius, it is just a fact that we do distinguish between hallucinations and normal cases of experience. Indeed, as we have seen, we usually detect these deceptions by seeing how much our particular experiences cohere with the empirical laws that regulate the whole of our experience. The assumption that all of our experience is actually nothing but a trick of some Evil Genius does not mean that the above criterion cannot be applied. Actually, it is *part* of the sceptical hypothesis that nothing would change in our 'experience' if nothing corresponded to our representations. Even if nothing corresponded to our representations, we would be drawing the distinction between bits of normal 'experience' and deviant cases. Since the latter do not cohere with the rest, we would take them as 'hallucinations,' obviously not realizing that they are simply deviant hallucinations within the general 'big hallucination' to which our experience actually amounts.

The fact that even the sceptic is committed to the lawfulness of what appears to us is obviously still not enough to refute the sceptic. All we have achieved is the concession that, if the entire external world were a trick of some Evil Genius, we would experience it with the same regu-

larities to which we are accustomed. But Descartes would readily concede that much, and Kant is perfectly aware of this. In fact, he seems to hold that his criterion is at best superfluous with regard to refuting the sceptic because empirical idealism is 'already' refuted through different considerations:

> Empirical idealism, and its mistaken questionings as to the objective reality of our outer perceptions, is already sufficiently refuted, when it has been shown that outer perception yields immediate proof of something actual in space, and that this space, although in itself only a mere form of representations, has objective reality in relation to all outer appearances, which also are nothing else than mere representations; and when it has likewise been shown that in the absence of perception even imagining and dreaming are not possible, and that our outer senses, as regards the data from which experience can arise, have therefore their actual corresponding objects in space. (A376–7)

Thus Kant's strategy for removing the spectre that our experience may be, *systematically* and not just on certain occasions, a mere product of the imagination, does not rest on the criterion introduced for detecting *particular* hallucinations. Rather, it rests on the following two claims, each of which Kant seems to take as sufficient to refute the sceptic: (a) the immediacy thesis, and (b) the thesis that the imagination is dependent on outer sense (and thus that the reality of the latter cannot be denied without also removing the very possibility of the former).[21] We have already seen why the immediacy thesis, although sound, by itself is not sufficient for the refutation. Can we then complete the argument by relying on the other thesis?

Unfortunately, as critics have largely recognized, the thesis of the imagination's dependence on outer sense is highly unsatisfactory.[22] It rests on a bold limitation of our imaginative power that, without clear justification, is assumed to be unable *by itself* to produce outer representations. Moreover, as Allison has pointed out, even if one concedes for the sake of argument that the faculty of imagination, *as we know it*, suffers this limitation, 'the possibility still remains that our representations of outer things are the results of some unknown "hidden faculty."'[23] Descartes himself mentions this possibility in the Third Meditation when he addresses the problem of the source of ideas that seem to come from outside us. According to Descartes, the observation that these ideas are independent of our will is not sufficient reason to consider

them as externally caused. He bases this claim on the following consideration: 'Perhaps there is in me some faculty or power adequate to produce these ideas without the aid of any external objects, even though it is not known to me.'[24] Finally, this line of thought seems to ignore the central idea that underpins the Cartesian hypothesis of the Evil Genius. In fact, Descartes's intention is precisely to raise the possibility of a power that goes well beyond the power of the imagination as we know it.[25] Thus Kant's strategy in the Fourth Paralogism suffers a serious difficulty. Its failure, at least as it stands, leads us to wonder whether the problem raised by the possibility of a 'super-imaginative power,' can be dealt with differently.

The Refutation of the Sceptic: The Alternative Strategy

I suggest that an alternative refutation can be mounted by combining two points, both of which depend on transcendental idealism: (1) the abandonment of the transcendental realist's picture of perception (the result achieved in the proof of the immediacy of outer perception), and (2) a reflection on the meaning, within a transcendental idealistic perspective, of the very possibility of a super-imaginative power that generates our *entire* experience, even in its lawfulness (which distinguishes the 'big hallucination' from occasional and thus non-problematic hallucinations).

To begin with, the sceptic assumes that both in the case of normal experience and in the case of hallucinations we perceive an 'idea' (with a spatial content), and that the difference between them is that in the latter case no object corresponds to the 'idea.' The immediacy thesis refutes this fundamental premise of the sceptical argument by establishing that whenever I perceive an external object, I do it immediately. As we saw, the existence of (particular) hallucinations is not sufficient grounds for inferring that what I normally see is an 'idea' in my mind. Now, my entire experience teaches me that the world is constituted by two distinct sets of objects: merely temporal inner states (and myself as the owner of these states), and spatio-temporal objects. Particular hallucinations certainly exist, but they are easily detectable through the above-mentioned criterion. Since they can be recognized only against the background of experience, their existence does not change in the least the general picture that my entire experience gives me of the world. It follows that when the sceptic asks the fatidic question, 'How do you know that what happens in the particular cases of hallucina-

tions is not the general rule?' we need not confess ignorance, as we would do if we accepted the assumption that all my experience concerns is 'ideas.' On the contrary – we can start replying (and this is crucial) that what my entire experience teaches me is that there are external things, mental entities, and hallucinations.

Now, to wonder whether the world is different from the one my entire experience informs me about obviously means to wonder how the world is from a perspective external to experience itself. The fact that we are talking about our *entire* experience or about *all* we know is crucial. When the sceptic asks us to wonder whether our entire experience could be illusory, he is really asking us to see the world from a point of view external to experience – that is, from an absolute standpoint. From the perspective of transcendental idealism, however, this means wondering how the world is in itself – namely, how it is from a point of view that abstracts from the way in which my sensibility makes things appear – that is, in space and time. But it is clear that if this is what the Cartesian sceptic is asking us, his challenge – the possibility that one's entire external experience is an illusion – turns out to be an illegitimate concern. It turns out to be a vain inquiry into the nature of the thing in itself.

We can reach the same result if we construe the sceptical challenge as raising the possibility that the world is a product of some super-imagination.[26] To say that spatial objects could be systematically just a product of an unknown faculty or super-imagination is to raise a question the answer to which *by definition* falls outside the sphere of possible experience. It means wondering about the nature of what affects my senses, before they organize the material resulting from the affection into a spatio-temporal object. But it is not this unknown X whose existence we need to prove in order to ground an empirical kind of realism. As Kant puts it: 'We can indeed admit that something, which may be (in the transcendental sense) outside us, is the cause of our intuitions, but this is not the object we are thinking in the representations of matter and of corporeal things' (A372). If we stay within the limits of possible experience and do not attempt to determine what matter (or the soul) in itself is, the question as to whether what we immediately perceive as an external, spatial world is different from the way it appears, becomes absurd: 'If, as the critical argument compels us to do, we hold fast to the rule above mentioned [do not take matter and the soul as things in themselves], we shall never dream of seeking to inform ourselves about the objects of the senses as they are in themselves, that is, out of all relation to the senses.'[27]

In other words, it is quite possible that the cause of the affection that provides all the material of my experience is some sort of transcendent Ego endowed with an unknown (super)imagination. This transcendent Ego would play precisely the same role as does Descartes's Evil Genius. The problem is that once again the possibility of raising such a scenario presupposes that one can appeal to how things are independently of the organization of my senses. In the present case, it presupposes the possibility of appealing to the unknown cause of our intuitions. In some revealing passages found in the quite different context of the Second Paralogism, Kant provides the most explicit expression of the refutation/dismissal of the radical doubt that we are proposing. He writes: 'The something that underlies the outer appearances and which so affects our sense that it obtains the representations of space matter, shape, etc., may yet, when viewed as noumenon (or better, as transcendental object), be at the same time, the subject of our thoughts' (A358). Also: 'I may further assume that the substance which in relation to our outer sense possesses extension is in itself the possessor of thoughts' (A359). And finally: 'What, as thing in itself, underlies the appearances of matter, perhaps after all may not be so heterogeneous in character [compared to what underlies the appearance of the soul]' (B428). And he emphasizes the very same point at the end of the Fourth Paralogism: 'Though the "I," as represented through inner sense in time, and objects outside me, are specifically quite distinct appearances, they are not for that reason thought as being different things. Neither the *transcendental object* which underlies outer appearances nor that which underlies inner intuition, is in itself either matter or a thinking being, but a ground (to us unknown) of the appearances which supply to us the empirical concept of the former as well as the latter mode of existence' (A379–80).

Kant's point in these passages is clear enough. If we look at a thing from an absolute standpoint and not from the only one that is given to us – that is, sensible experience – then we can quite naturally concede to the empirical idealist that what appears as extended matter is 'in itself' the same thing as the subject – or, more precisely, merely an inner state of the (transcendental) subject. From such an absolute viewpoint, any kind of metaphysical speculation is allowed. Matter may be spirit in itself, or spirit may be matter in itself, or both matter in itself and spirit in itself may exist:

> If the psychologist takes appearances for things in themselves, and as existing in and by themselves, then whether he be a materialist who

> admits into his system nothing but matter alone, or a spiritualist who admits only thinking beings (that is, beings with the form of our inner sense), or a dualist who accepts both, he will always, owing to this misunderstanding, be entangled in pseudo-rational speculations as to how that which is not a thing in itself, but only the appearance of a thing in general, can exist by itself.[28] (A380)

We are now in a position to see that once we have abandoned the idea that the immediate object of our knowledge is a mental entity, the hypothesis of an unknown super-imaginative power is equivalent to the possibility that the non-sensible cause of our representations – that is, the non-sensible correlate of appearances – is this imaginative super-power. This, however, simply gives a new name to what Kant calls an external (in the transcendental sense) cause of our representations. The hypothesis of the super-imagination is threatening for a transcendental realist (it would remove the objects that our knowledge is supposed to be about); yet the same hypothesis does not concern the transcendental idealist because it would boil down to substituting the term 'non-sensible cause of our representation' with the more suggestive 'super-power of imagination' or 'Evil Genius.' Once this super-imagination is shown to be necessarily confined to a non-sensible sphere, it becomes incapable of threatening empirical realism. We thus understand Kant's crucial remark that 'even the most rigid idealist cannot, therefore, require a proof that the object outside us (taking 'outside' in the strict [transcendental] sense) corresponds to our perception. For if there be such an object, it could not be represented and intuited as outside us' (A375–6).

Note how the argument I have just presented rests on a complete reversal of Kant's refutation in the *Dissertation*. While the 1770 argument turned on the affirmation of the genuine external existence of things in themselves, much of the present argument turns precisely on the possibility that things in themselves are not external and 'real,' but rather similar to some ideal cause such as the Evil Genius or an unknown imaginative power. Conversely, while the 1770 argument conceded to the sceptic that phenomena are merely mental entities, the present argument relies on the ontological mind-independence of phenomena. While the 1770 dogmatic interpretation of the things in themselves/phenomena distinction was the greatest obstacle put in Kant's way to refute the sceptic, the mature, truly *critical* interpretation of the same distinction constitutes the necessary background from which the new argument arises.

A Crucial Objection Answered

It should be clear by now that Kant's reply to the sceptic turns on reinterpreting the Evil Genius hypothesis in terms of a question that, systematically surpassing the *totality* of any possible experience, brings us into the field of the thing in itself. The success of this argument rests precisely on the legitimacy of this reinterpretation. One could object that Descartes's hypothesis arises from familiar cases of delusions of the imagination, such as hallucinations. As such, it certainly does not exceed the limits of ordinary experience. The sceptic needs only to affirm that unproblematic familiar cases of hallucinations could simply be the general rule. More precisely, particular hallucinations give us the opportunity to conceive of the possibility that – in analogy to what happens with these particular cases of delusion – our entire experience is a sort of 'big hallucination,' in which particular hallucinations would simply be deviant cases of the regular hallucinatory 'experience' in which we live. And it is not clear how, reasoning in this manner, the skeptic could arrive at the realm of the thing in itself. The possibility that he raises does not seem to have anything to do with the constitution of the noumenal world. Even if we assume transcendental idealism, it still seems that the possibility raised by the sceptic is perfectly legitimate. Doesn't the transcendental idealist acknowledge that hallucinations exist? How then can he deny the legitimacy of the hypothesis that our entire experience is nothing but a systematic hallucination?

At this point the immediacy thesis becomes crucial. As we saw, if we operate with the transcendental realist's model of perception, then the objects we experience (mental entities) are not those which we want to reach (objects external in the transcendental sense). Therefore, when we refer to our entire experience, we are always talking about a mental realm. It follows that the Cartesian hypothesis does not even become a question that surpasses the limits of our experience: it is simply a question about the possibility that our entire experience could lack its desired referent. But if we abandon the transcendental realist's model of perception – that is, if we accept the immediacy thesis – the only way to make sense of the idea that my entire external experience is a 'big hallucination' is to reject this experience and raise the possibility that the world, from a perspective other than the empirical one, could be different. In other words, the sceptical question can no longer be, 'Is there a correspondence between my entire experience (mental entities) and the desired objects?' The question *must* become, 'Could my entire

experience, that is, the objects that I immediately perceive, be different from the way they appear?' This is, however, equivalent to saying that things in themselves could be (or simply are) different from the way in which the empirical objects appear. Strictly speaking, it is not that Kant could readily accept this, nor even that he could simply say we know nothing about this realm. It is rather that he would dismiss the question itself as illegitimate, exceeding the set of questions we can raise. As Kant puts it very clearly in a footnote in the Transcendental Dialectic: 'Although to the question, what is the constitution of a transcendental object, no answer can be given stating *what it is*, we can yet reply that the *question* itself is *nothing*, because there is no given object [corresponding] to it' (A479/B507n).

Clearly, this response is possible only from the standpoint of transcendental idealism because only from that standpoint does the question about the constitution of the transcendental object amount to nothing. To be sure, other philosophical standpoints can reject the sceptical hypothesis on the basis of the generic charge that it rests on a scenario that goes systematically beyond our experience. This is the case with Carnap's philosophy, which attempted to refute idealism precisely along these lines. Despite superficial similarities to Kant's refutation, however, Carnap's approach is profoundly different and open to serious difficulties precisely because it lacks a transcendental idealistic perspective. The analysis of Carnap's refutation – which we assume as paradigmatic of all refutations turning on a generic appeal to test within the limits of possible experience – should therefore help us see the force of Kant's idea that transcendental idealism is 'the only refuge.'

Transcendental Idealism as 'the Only Refuge'

Carnap dismisses idealism because in his view it is grounded in a metaphysical and therefore meaningless perspective. For Carnap, a question or a sentence is meaningful if and only if it describes a state of affairs that can be verified or falsified through experience. Since scepticism raises the possibility of a state of affairs (the world as a systematic hallucination) that *ex hypothesi* cannot be verified or falsified through experience, it is a meaningless position. Carnap's emphasis that we must remain within the domain of experience resembles the kind of Kantian argument we have proposed. But this resemblance has limits, and it is precisely the determination of these limits that will display at the same

time the importance of transcendental idealism for our argument and its superiority over other refutations of the sceptic.

No matter how liberalized, the empirical verifiability principle of meaningfulness makes (as Carnap well knows) not only idealism but also realism meaningless. The issue between a realist and an idealist regarding the existence of any particular thing is merely a pseudo-dispute. The realist argues that the object exists independently of its being perceived, and the idealist argues for the coincidence of its existence and its being perceived. Thus 'there is complete unanimity so far as the empirical facts are concerned,'[29] as in principle there is no empirical fact that could settle the dispute between them. Moreover, what is true of the dispute over a particular object is true of the world in general. It follows that the dispute between idealism and realism – the world exists outside our mind versus the world is (or may be) identical with its being perceived – is a pseudo-issue as well.

Carnap's idea that between the idealist and the realist 'there is complete unanimity so far as the empirical facts are concerned' is highly instructive. It is not clear at all that by claiming that 'there is the object X out there,' a realist means the same thing of an idealist; the only way in which there could be unanimity is if the realist assumes that the immediate objects of his perception are mental entities. These are the 'empirical facts' (about which the idealist and the realist agree) that Carnap seems to have in mind. The disagreement between the idealist and the realist is merely over whether there is a corresponding object that exists independently of the mind. Thus the meaningless realist affirmation is that there is such an object, whereas the equally meaningless idealist affirmation is that there is no such object.

For Carnap, that we speak of external objects and therefore assume that they exist is merely the outcome of the fact that we have chosen a language that includes terms referring to abiding external objects. We choose this language simply because it is more efficient than a phenomenological language that includes only sense-data. But this efficiency is not to be taken as evidence in favour of the existence of external things. For Carnap, it is a merely practical advantage: it makes the verification or falsification of the sentences we utter easier. More precisely, a language that includes external things seems more efficient than any other for organizing and expressing our sense-data in such a way that our sentences can be verified or falsified. Thus the sense-data have priority.

Clearly, from a Kantian perspective Carnap is operating within a transcendental realist model. The immediate objects of our experience

are mental entities. The only difference between the transcendental realist and Carnap is that the former wants to affirm the correspondence of objects to these mental entities as a matter of fact, independent of the language we choose, whereas Carnap judges the question as meaningful only within a certain language – specifically, the one that includes external things. It is also clear that the sceptic cannot be threatened by such an attempt to remove the meaningfulness of his position. The sceptic can be satisfied with the fact that on this account, his realist opponent also ends up in the same spot (his position turns out equally meaningless). He can even consider Carnap an ally, notwithstanding his tendency to reduce idealism to non-sense. Indeed, the freedom that Carnap gives us to choose between a language that includes external things and one that restricts itself to sense-data is really equivalent to the sceptical idealist's thesis that there is no way to determine whether a world corresponds to our sense-data. After all, that the realist and idealist positions are really indistinguishable is precisely what the *sceptical* idealist wants. Indeed, if there is a version of idealism that Carnap can be legitimately said to oppose, it is what Kant calls *dogmatic* idealism. The ability of Carnap's position to undermine sceptical idealism is another matter. Carnap's implicit acceptance of the priority of sense-data and his insistence on the merely practical nature of the choice of the language through which we express these data make his position strikingly similar to sceptical idealism.

Similar considerations apply to the case of contemporary versions of verificationism, such as that of Dummett. Notoriously, Dummett advocates the view that there is no truth (in any domain) *without evidence or at least the possibility of verification*. Unlike the early Carnap, however, Dummett denies that statements which are unverifiable, even in principle, are ipso facto meaningless. These statements, in his view, have meaning but possess no truth value. Since statements that at a certain point in time lack any direct or indirect evidence might later be attributed some evidential basis or even be proved, Dummett believes that certain statements are neither true nor false, but simply 'not true.' The class of the 'not true' statements is the class of propositions that lack evidence for their truth or their falsity. The proposition 'π is a transcendental number,' for example, was 'not true' until 1882, but became true when Ferdinand von Lindemann proved it.[30] For a realist, quite to the contrary, this proposition has been true all along and was 'merely' proved in 1882. This explains why Dummett rejects what he calls the principle of bivalence (at times considered equivalent to that of the

excluded middle), whereby a proposition is always either true or false, and why he considers the acceptance of the same principle as what is 'integral' (read 'essential') to realism.

From this perspective, one might be tempted to argue that it is possible to mount a refutation of scepticism similar to the Kantian argument we have advocated. The argument would be roughly this: Scepticism, as we said, asks a question that systematically escapes any possible experience. It is difficult to imagine an experiential event that would falsify (or confirm) the thesis that at least one experience of external objects is not hallucinatory. While we can never know whether this particular experience is a hallucination, at the same time it is impossible to know whether this same experience is a trustworthy experience of a genuine external object. Hence, either there is a rational (as opposed to empirical) refutation of scepticism, such as Putnam's Brains in the Vat argument (on this more shortly), or scepticism must fall in the 'not true' category.

Now, in a Dummettian context, if there were such a rational argument, scepticism would no longer be 'not true' but simply false, and realism about the external world would move from the 'not true' category to the 'true' category. To find such an argument, however, would be to ground one kind of realism (the one regarding the material world) that Dummett apparently intends to attack. After all, his only quarrel with the phenomenalists seems to be their absent-minded acceptance of the principle of bivalence combined with their unnecessary attachment to strong reductionism (the translatability of a language that contains tables and chairs into a language of sense data). Therefore, Dummett cannot concede even the possibility of such a rational refutation. Now, if neither an experiential nor a rational refutation of scepticism is possible within Dummett's framework, we are perhaps authorized to conclude that the inclusion in the 'not true' category is precisely Dummett's way of dealing with scepticism.

There is, no doubt, a superficial resemblance between this line of thought and our argument. Both attempt to silence the sceptic, not through a direct refutation, but through a reconsideration of the very meaning of the sceptical challenge. Two things, however, make this argument essentially different from Kant's. To begin with, as we saw, realism about the external world turns out to be as 'not true' as scepticism is and for the very same reasons. Second, and consequently, scepticism on Dummett's reconstruction is by no means 'illegitimate'; again, it is simply 'not true' (like empirical realism). Therefore, to tackle

the question of scepticism from a Dummettian perspective means to aim, at most, to show that scepticism and antiscepticism about the external world are on the same footing. While this seems to introduce a difference with Carnap – whose 'refutation' of scepticism paradoxically turned out to be a confirmation of dogmatic idealism – it is clearly insufficient for grounding that kind of independent existence of the external world which Kant (and any adversary of scepticism) was after. This result is hardly surprising: properly understood, in a Kantian context knowledge is about an ontologically independent realm (yet conditioned by our epistemic forms) – that is, the realm of appearances. The question of whether there is at least one non-hallucinatory experience has a very clear answer: yes. All experience that does not seem to violate the 'rules of experience' is about an objective realm that exists independently of the fact whether someone experiences it or not (although, obviously, *how* one experiences it depends on subjective forms). In Dummett, any reference to this independent realm would collide with his modified verificationism. Knowledge is not about an independent realm; rather, if the provocation is allowed, it is about our experience. From this perspective, any argument against scepticism can at most hope to tie the match with it.

The result of this analysis is that the mere appeal to the need to remain within the domain of experience is not sufficient to silence the sceptic and should not be confused with our antisceptical argument. If the idea that outer perception is immediate is not brought in, the mere appeal to what experience says is never going to be sufficient. Only if this appeal is combined with the immediacy thesis (and therefore with transcendental idealism) can scepticism be refuted. These were, in fact, just the two steps of our argument. We have already seen how the thesis of the immediacy of outer perception is by itself not sufficient. We now learn that the second step as well – at least, if it is taken as a generic suggestion to remain within the limits of experience – is by itself not sufficient. The appeal to experience as the framework within which questions are legitimate is not going to yield the desired result unless the experience we are referring to is sharply distinguished from the sense-data that Carnap assumes, or from a realm whose ontological independence is not affirmed by the epistemological framework we use, as in the case of Dummett. Only if we mean by 'experience' the world we immediately see, the spatial world given to us independently of the nature of the supersensible thing that affects us, can we really reinterpret and rule out the sceptical hypothesis as a illegitimate question. The idea of immedi-

acy can be convincingly grounded only from a transcendentally idealist perspective, as we saw in our discussion of Reid. It follows that once again, transcendental idealism turns out to be the crucial and necessary premise of a successful refutation of the sceptic – or to repeat our favourite expression, the 'only refuge.'[31]

Kant and Putnam

The necessary role we have just assigned to transcendental idealism could be further questioned by referring to other apparently successful refutations that arise from a broadly construed Kantian framework, but that seem to bypass any appeal to idealistic premises. This is the case with Putnam's famous argument against the possibility that we are 'brains in a vat.' Putnam rephrases Descartes's hypothesis of the Evil Genius by imagining that we are brains placed in a vat containing a nutrient fluid and connected through a series of wires to a sort of supercomputer – the Matrix, we could say, after Hollywood's appropriation of this philosophical fantasy. The question obviously is this: How do we know that we are *not* brains in a vat?

Putnam thinks that a bit of reflection on how words refer to objects suffices to show that this scenario is a mere *logical* possibility, not a *real* one. To put it differently, it is a merely *consistent* story, but it cannot be the description of how things actually are. In particular, Putnam believes he can show that saying 'we are brains in a vat' is no different from saying 'I do not exist' or 'all general statements are false.' These two assertions are evidently self-refuting (albeit for different reasons). One approach to Putnam's argument is to see it as an attempt to show that the sentence 'we are brains in a vat,' whose inner logic does not seem to share any similarity with the two sentences just cited, turns out to be equally self-refuting. In other words, if Putnam's argument succeeds, then the sceptical position is refuted in a very straightforward way: it is built on a basic contradiction.

The argument goes as follows. Let us assume that we are in fact brains in a vat. When a brain in a vat thinks 'there is a tree in front of me,' he is not referring to a real tree (an external object as we understand it); by definition, if his thought refers at all, it refers to a 'tree in the image' – that is, to the image stimulated by the electrical impulses sent by the supercomputer. Something similar happens for any affirmation the brain might make about external objects, including the brain itself, understood as an external object in a way not different from the

way in which we take our body as an external object. It follows that if the brain thinks (or says) 'I am a brain in a vat,' then it *must* mean 'I am a brain in a vat in the image.' Therefore, if we *are* brains in a vat, when we say 'we are brains in a vat', we are, so to speak, condemned to mean that we are 'brains in a vat in the image.' It was our intention to say something *not* about our internal sense data; yet the logic of our language – given the presupposition that we *are* brains in a vat – denies the possibility of reaching real things beyond them. This means, however, that if we are brains in a vat, then our sentence 'we are brains in a vat,' being part of a language that is forced to refer to inner sense data, affirms something false. Now, if the truth of a state of affairs described by a sentence removes the conditions of possibility of the sentence's being true, then the sentence is necessarily false. To quote Putnam's succinct conclusion to his argument: 'If we are brains in a vat, then "We are brains in a vat" is false. So it is (necessarily) false.'[32]

The 'vat-English' spoken by the brains has no chance of going beyond the level of sense data to which, by the very setting up of the experiment, those brains are condemned to remain confined. The rules that govern reference in vat-English exclude the possibility of reaching external objects. And when it comes to describing those few objects that constitute the poor furniture of the universe in the sceptical hypothesis, the self-refuting nature of the hypothesis itself becomes evident. This explains the self-refuting status of the affirmation 'we are brains in a vat.' A reflection of the conditions of the possibility of reference – as opposed to an appeal to transcendental idealism – has yielded a successful refutation.

In this ingenious argument, is there any appeal to transcendental idealism? Apparently not. All we have been given is a reflection on the logic of reference and on how this leads to the self-refutation of the sceptic. Note also that the impression that Putnam's argument is completely independent of transcendental idealism is strengthened by the fact that his argument is explicitly meant to be used, as an antisceptical tool, by what he calls the 'externalist' philosopher. In other words, the 'externalist' philosopher should be able to appeal to it in order to quash the sceptical challenge. Since such a philosopher believes that truth consists of correspondence between words (or thought signs) and objects *as completely independent of the system of description used* (in our language, completely independent of the mind), he is in fundamental disagreement with the transcendental idealist. In fact, what Putnam calls the 'externalist philosopher' is very close to what Kant calls the 'transcendental

realist.' Now, if Putnam's argument is to be used by the externalist philosopher, then clearly no internalist (let alone idealist) premise can be assumed in the argument. In fact, Putnam believes he has shown that 'the very relation of correspondence on which truth and reference depend (on his [*the externalist's*] view) cannot logically be available to him if he *is* a Brain in a Vat. So if we *are* Brains in a Vat, we cannot *think* that we are, except in the bracketed sense [we are Brains in a Vat]; and this bracketed thought does not have reference conditions that would make it *true*. So it is not possible after all that we are Brains in a Vat.'

In other words, Putnam assumes the externalist perspective and shows that, given the theory of reference that follows from that perspective, the Brains in a Vat hypothesis is self-refuting. In this sense, his argument is primarily meant to be a tool at the disposal of the externalist philosopher to rebut the sceptic. Note that from the internalist perspective, this ingenious argument need not even turn on the conditions of the possibility of reference. If we assume Putnam's internal realism (or Kant's transcendental idealism), then the Brains in a Vat hypothesis becomes 'just a story' – a mere *logical* possibility, to use Kant's language. In fact, the sceptical story, as we have shown, implicitly assumes a God's eye point of view; it is told from the perspective of an 'eye' that does not belong to the world it sees. But to repeat, Putnam believes that his argument is capable of refuting the sceptic *even if one assumes the externalist perspective*; in this way he leaves aside any considerations regarding the opportunity to avoid a God's eye point of view.

Now, if this is true, the *need* to assume transcendental idealism in order to avoid the sceptical threat – a need on which we have so much insisted – would be readily falsified. In other words, if Putnam has correctly assessed the logic of his own argument, then transcendental idealism is *not*, as Kant argues, 'the only refuge.' On closer analysis, however, this reconstruction proves to be mistaken. Let us return to Putnam's argument. Its main thrust is that if we are brains in a vat, we cannot *think* (or *say*) that we are. More precisely, if we are brains in a vat and (therefore) we have to speak vat-English, we fall into self-contradiction as soon as we try to express our 'true' condition (*being* brains in a vat). It is rather evident, though, that this leaves intact the possibility that we *are* indeed brains in a vat *from the perspective of a putative external observer* (be it God or any other entity that falls outside the world we inhabit). The argument only shows that from within our world, we cannot even *express* our wretched condition. *From our perspective*, the Brains in a Vat hypothesis is a mere logical possibility, because it violates the

conditions of possibility of reference. But the sceptic can very well reply that this is perfectly compatible with the possibility that the whole of our knowledge, if seen from the perspective of God, amounts to a systematic error, with the further complication that we would not even be able to formulate – without falling in contradiction or paradox – why this is so (that is, why we are brains in a vat). In sum, it is still possible that we are brains in a vat, that the externalist view of truth is right, and that truth (so defined) can never be achieved – not even in the case of the description of why this state of affairs holds.

Now, saying that this possibility does not concern us because it presupposes a point of view external to the world we inhabit amounts to assuming already some forms of internalism (or transcendental idealism).[33] Certainly, Putnam might reply that our objection presupposes, not just an 'external viewpoint,' but a viewpoint that transcends even the conditions of possibility of reference of our language. This reply, however, does not change much. Why, in fact, should this external viewer be bound to such conditions? Couldn't he be bound to reference conditions (if any) completely different from ours? And if so, wouldn't he be completely safe from the paradox that Putnam has shown to be an inevitable destiny that *we* face in describing our condition? Thus, in order to be satisfied with the fact that, if we are brains in a vat, then we cannot express this condition, we need to have already sided with a philosophical perspective that wants to remain within the sphere of successful (that is, non-paradoxical) use of our language, something very close to some form of internalism.[34]

From this, four main consequences follow. First, despite appearances, transcendental idealism (or internalism – the difference here does not matter) *is* a necessary premise of Putnam's antisceptical argument. Second, Putnam, unlike Kant, fails to make this logical dependence clear. Third, insofar as Putnam interprets his argument as a foundation of internalism, he seems to be arguing in a circular fashion.[35] And fourth, and most importantly for our purposes, Kant's argument, as we have reconstructed it, remains on a firmer footing than Putnam's, simply because its main premise (the truth of transcendental idealism) is not only clearly acknowledged, but *grounded* through the arguments we analysed in the third chapter, while I fail to see where Putnam's internalism finds a comparable foundation.

according to Putnam, is straightforward: the Brains in a Vat hypothesis cannot even be raised within a perspective that denies the legitimacy of a God's eye point of view – that is, the point of view from which the 'story' is necessarily told.

4. The Problem of Idealism between 1781 and 1787

1 I. Kant, *Prolegomena* (1977), 100.
2 Ibid., 30.
3 Ibid., 72.
4 I. Kant, Ak. 4: 376.
5 I. Kant, *Prolegomena* (1977), 107.
6 Ibid.
7 The importance of the idealism issue in these years is also revealed in the exchange of letters between Kant and Garve following the 1782 review. For a complete analysis of this exchange see Heidemann, *Kant*, 87–94.
8 See Ak. 29: 928–9.
9 Kant, *Prolegomena*, 43.
10 See, *La doctrine*, 146–7.
11 Rousset refers to Kant, *Prolegomena*, 43.
12 Kant writes: 'The empirical truth of appearances in space and time is, however, sufficiently secured; it is adequately distinguished from dreams, if both dreams and genuine appearances cohere truly and completely in one experience, in accordance with empirical laws' (A492/B520).
13 See A493/B521. These laws of the empirical advance are simply the transcendental laws that make the perception of an event possible (for example, necessary presence of a cause for a given effect) and that are therefore supposed to be respected also in the perception of new events – namely, in the empirical advance.
14 Most importantly, the Garve-Feder review (see *Zugabe zu den Göttingischen Anzeigen von gelehrten Sachen*, Göttingen, 1770–82, III, 40–8). This review and the polemical discussion that followed between Kant and Feder after 1781 has been shown by Heidemann to have been extremely important for the Second Edition of the *Critique* – specifically, for the removal of the Fourth Paralogism and the elaboration of a new refutation of idealism (Heidemann, *Kant*, 87–94).
15 That Kant's intentions were not to move from phenomenalism to non-phenomenalism in the passage from the Fourth Paralogism to the Refutation of Idealism is clearly indicated by Kant himself when he notes that the new

proof differentiates itself from the original for 'the method of proof only'(Bxl n). For a completely different assessment of this remark, see Guyer, *Kant*, 281–2. Guyer, influenced by his phenomenalistic reading of the Fourth Paralogism, tends to interpret this remark as misleading because it would suggest that the 1787 proof itself is infected by phenomenalism, whereas for him, one of the great advantages of the Refutation is precisely that Kant there happily combines 'epistemological subjectivism' with 'ontological realism' – that is, the view that objects as we know them are different from the way they are in themselves, but that they are nonetheless numerically distinct from the subject.

16 The idea that Kant removed the Fourth Paralogism because he no longer viewed it as a viable refutation is often a consequence of the assumption that the Fourth Paralogism is committed to phenomenalism. The clearest expression of this view is perhaps to be found in Rousset (see *La doctrine*, 148).

17 See Bxl n and B275.

18 See Bxl.

19 Ibid.

5. The Refutation of Idealism

1 In the first Note that Kant inserts after the proof, he criticizes scepticism because it assumes that 'the only immediate experience is inner experience, and that from it we can only *infer* outer things.' This suggests that the 1781 immediacy thesis was still being endorsed by Kant in 1787. Note however that, as Kant says in the footnote attached to the Note, the immediacy of outer experience is not assumed (as in the Fourth Paralogism, where it was simply introduced as a consequence of transcendental idealism); rather, it is proved. Indeed, the proof allegedly shows that inner experience is possible only on the condition that we have experience of things, not of representations thereof, which means that we have immediate outer experience.

2 See A218/B266. This point was also made in the Fourth Paralogism: 'It is sensation, therefore, that indicates a reality in space and time' (A373).

3 B277n.

4 A. Zweig, (ed.), *Philosophical Correspondence*, 61 (Ak. X: 102).

5 See B157, B158n, B422n.

6 See A33/B49–50, A34/B50–1, and B67.

7 In 'Personal Identity and Kant's "Refutation of Idealism"' (259–78), Aquila