

Routledge Philosophy GuideBook to

Kant and the *Critique
of Pure Reason*

▪ *Sebastian Gardner*



LONDON AND NEW YORK

Contents

	<i>Preface</i>	viii
1	The problem of metaphysics	1
2	The possibility of objects	18
3	How are synthetic a priori judgements possible? (The Introduction)	34
4	The sensible conditions of objects (The Aesthetic)	43
5	Transcendental idealism	57
6	The conceptual conditions of objects (The Analytic)	75
7	Unknowable objects (The Dialectic)	136
8	The meaning of transcendental idealism	175
9	The complete Critical system (The Canon of Pure Reason)	199
10	The reception and influence of the <i>Critique</i>	212
	<i>Bibliography</i>	226
	<i>Index</i>	236

Chapter 2

The possibility of objects

After a brief summary of the Copernican revolution in the Preface, where it is advanced as the general solution to the problem of metaphysics, Kant leaves it to the reader to extrapolate the exact nature of his philosophical revolution from the specific doctrines that follow in the *Critique*. Understanding the arguments of the *Critique* is however made considerably easier by having in advance a fuller idea of the Copernican revolution than can be gleaned from the Preface. Accordingly, this chapter attempts to set out the line of thought underlying Kant's Copernicanism, in terms that as far as possible avoid the technicalities of his philosophy.

The Critical problem: Kant's letter to Herz

The place to start is with the letter referred to earlier in which Kant acknowledges the failure of the *Dissertation*, and first states the Critical problem. Kant is talking about a work that he had previously planned with the projected title, 'The limits of sense and reason':

As I thought through the theoretical part, considering its whole scope and the reciprocal relations of all its parts, I noticed that I still lacked something essential, something that in my long metaphysical studies I, as well as others, had failed to pay attention to and that, in fact, constitutes the key to the whole secret of hitherto obscure metaphysics. I asked myself: What is the ground of the relation of that in us which we call 'representation' to the object? If a representation is only a way in which the subject is affected by the object, then it is easy to see how the representation is in conformity with this object, namely, as an effect in accord with its cause, and it is easy to see how this modification of our mind can represent something, that is, have an object. Thus the passive or sensuous representations have an understandable relationship to objects, and the principles that are derived from the nature of our soul have an understandable validity for all things insofar as those things are supposed to be objects of the senses. In the same way, if that in us which we call 'representation' were active with regard to the object, that is, if the object itself were created by the representation (as when divine cognitions are conceived as the archetypes of all things), the conformity of these representations to their objects could be understood. Thus the possibility of both an *intellectus archetypi* (on whose intuitions the things themselves would be grounded) and an *intellectus ectypi* (which would derive the data

for its logical procedure from the sensuous intuitions of things) is at least intelligible. However, our understanding, through its representations, is not the cause of the object (save in the case of moral ends), nor is the object the cause of the intellectual representations in the mind (*in sensu reali*). Therefore the pure concepts of the understanding must not be abstracted from sense perceptions, nor must they express the receptions of representations through the senses; but though they must have their origin in the nature of the soul, they are neither caused by the object nor bring the object itself into being. In my dissertation I was content to explain the nature of intellectual representations in a merely negative way, namely, to state that they were not modifications of the soul brought about by the object. However, I silently passed over the further question of how a representation that refers to an object without being in any way affected by it can be possible. I had said: The sensuous representations present things as they appear, the intellectual presentations present them as they are. But by what means are these things given to us, if not by the way in which they affect us? And if such intellectual representations depend on our inner activity, whence comes the agreement they are supposed to have with objects – how do they agree with these objects, since the agreement has not been reached with the aid of experience? . . .

Plato assumed a previous intuition of divinity as the primary source of the pure concepts of the understanding and of first principles. Malebranche believed in a still-continuing perennial intuition of this primary being. Various moralists have accepted precisely this view with respect to basic moral laws. Crusius believed in certain implanted rules for the purpose of forming judgements and ready-made concepts that God implanted in the human soul just as they had to be in order to harmonize with things. Of these systems, one may call the former [Plato and Malebranche] *influxum hyperphysicum* and the latter [Crusius; Kant might also have referred to Leibniz] *harmonium praestabilitam intellectualem*. But the *deus ex machina* is the greatest absurdity one could hit upon in the determination of the origin and validity of our knowledge. It has – besides its deceptive circle in the conclusion concerning our cognitions – also this additional disadvantage: it encourages all sorts of wild notions and every pious and speculative brainstorm.

(Letter to Herz, 21 February 1772)

Representation (*Vorstellung*) is Kant's generic term for a constituent or element of cognition, similar in scope to 'idea' in the writings of the rationalists and empiricists: anything subjective that can play a role in composing a judgement or knowledge claim counts as a representation for Kant (its sense is thus philosophical, not psychological). It is in general possible to understand how a representation can relate to its object, Kant claims, if it either causes its object, or is caused by its object. Consequently there is no problem in understanding how sensory representations can relate to objects, since the mind is straightforwardly passive with respect to the objects of the senses, which (we naturally suppose) produce representations of themselves in us. But it is not readily

intelligible how intellectual representations – ‘the pure concepts of the understanding’, which the *Dissertation* had claimed can alone represent ‘things *as they are*’ – can relate to objects, for they are not produced through our being affected by objects (Kant assumes that empiricist accounts of concept-formation are false), and nor do they produce their objects (to suppose which would be to confuse the human intellect with the creative intellect of God). Hence the problem. Previous solutions to it, Kant points out, are evidently defective.

In the letter, Kant nevertheless goes on to claim that he possesses the solution to the problem, and says that he will publish ‘within three months’ a work beating the title ‘Critique of Pure Reason’. In fact the Critical problem of the ground of the relation of representation and object preoccupied Kant for the best part of a decade, and in the course of attempting to solve it, his view of it changed in two important respects: he ceased to think that a causal relation running from the object to the subject suffices to make it intelligible that the subject is able to represent objects; and, as a result, he came to think that the problem of agreement with objects is not restricted to intellectual representations but rather extends to all of our representations. The Critical problem which motivates the Copernican revolution in the *Critique* is thus a broader and deeper version of the problem identified in the letter to Herz. Furthermore, Kant discovered that the Critical problem does not admit of a solution under the rationalist assumption that he still upheld in the letter to Herz: there is, the *Critique* tells us, an answer to the question of how pure intellectual representations can apply to things ‘*as they appear*’, but not to the question of how they may be applied to things ‘*as they are*’.

Interpretations of Kant: analytic and idealist

This is the appropriate point at which to introduce an important distinction between two different ways of reading Kant’s strategy in the *Critique*. These may be called the analytic and idealist interpretations. Both are found in present-day English-language commentary. The following quotations give an idea of each:

It is possible to imagine kinds of world very different from the world as we know it. It is possible to describe types of experience very different from the experience we actually have. But not any purported and grammatically permissible description of a possible kind of experience would be a truly intelligible description. There are limits to what we can conceive of, or make intelligible to ourselves, as a possible general structure of experience. The investigation of these limits, the investigation of the set of ideas which forms the limiting framework of all our thought about the world and experience of the world, is, evidently, an important and interesting philosophical undertaking. No philosopher has made a more strenuous attempt on it than Kant.

(P. F. Strawson)

Such an account [of the constitution of the world] requires reference to the operations of the mind, without which the world in question would not be

disclosed to us and could not possibly adopt its shape. In this way Kant explained nature and the world of nature by means of rules that guide the synthetic activities we must exert on what is given to us in sensation.

But the source from which a world originates is equally dependent on that world. Initially it might seem that the principle by which we are capable of accounting for a world remains independent of what it accounts for. Closer investigation, however, discovers that, unless it executes the activities from which a world originates, the principle itself would be incomprehensible. This kind of investigation is distinctive to the method of Kant's epistemology that he calls 'transcendental': it can be shown that the unity of self-consciousness could not even be conceived unless that very unity functions as the point of departure for constituting a world of objects. With this, we can understand not only the origin of this world but also why this world is natural and indispensable to us and why our knowledge claims about it are justified.

(Dieter Henrich)

The analytic interpretation, represented by Strawson, has its name because it identifies the task of Kantian philosophy as that of analysing the implications of our conception of experience. It seeks to find in Kant, or reconstruct from materials supplied by him, what are known as transcendental arguments. These attempt to demonstrate that experience necessarily has certain features, ones which accord with commonsense realism. To take a central example of Strawson's, the *Critique* may be held to contain an argument to the effect that we could not conceive ourselves as subjects of experience if we did not have experience of a world of spatio-temporal particulars existing independently of our experiences. The chief point of uncovering the structure of experience, on the analytic interpretation, is that it allows skepticism to be refuted. The analytic line of interpretation thus construes Kant as employing a novel means, transcendental argumentation, in pursuit of the traditional epistemological goal of justifying our knowledge claims, and holds that what is of value in Kant's philosophy has to do with transcendental arguments, everything in it that comes under the heading of the Copernican revolution falling by the wayside.

The idealist line of interpretation, represented by the quotation from Henrich, agrees of course that Kant meant to provide a justification of our knowledge claims, but holds that he rightly intended this to follow from a more fundamental investigation into how the subject constitutes the world. This primary investigation of Kant's, on the idealist view, proceeds at a different level from traditional epistemology. The idealist interpretation agrees that Kant's enquiry is directed to uncovering the structure of experience but understands this notion differently. On the analytic interpretation, the structure of experience ultimately reduces to the structure of *what* is experienced: to say that experience has structure is just to say that it is necessarily *of* certain kinds of things (such as objective spatio-temporal particulars). On the idealist interpretation, experience *itself*, the activity of experiencing, has an inherent structure, which it bestows on its objects. Transcendental investigation shows that the operations of the mind give shape to the world, as Henrich puts it.

The two lines of interpretation thus disagree about the kind of philosophical explanation to be looked for in Kant. The analytic interpretation regards statements about the conceptual presuppositions of experience as self-sufficient, and the Critical problem as solved once the structure of experience has been specified. It grounds all claims about the structure of experience on an appeal to the impossibility of our forming any other conception of experience. The structure of experience, it holds, is nothing more than the necessary window onto the world, and cannot be said to give shape to it: that experience has a structure is ultimately just a matter of our having such and such concepts and being unable to conceive any alternative to them, and the attempt to invest it with metaphysical significance over and above the completely minimal sense of being necessary for experience it regards as gratuitous and erroneous.

The idealist interpretation, by contrast, sees the need for further explanation of the structure of experience, and it refers this structure to the operations of our mind. Why it takes this view, and why it considers that Kant was right to claim that the solution of the Critical problem requires the Copernican revolution, is explained in what follows.

The problem of reality

The deepest issue with which Kant is preoccupied, on the idealist interpretation, is that of the possibility of objects. To approach the special sense in which Kant regards the possibility of objects as standing in need of explanation, it is necessary to consider Kant's view of what may be called the problem of reality. This problem is a generalised version of the Critical problem identified in the letter to Herz.

There is, we naturally suppose, a real world. The proposition that there is such a thing as reality is one that can scarcely allow itself to be doubted. We suppose, furthermore, that reality is known or in principle knowable to us, if only in part. Reality is then naturally conceived as that which fundamentally explains how objects of experience and thought are possible for us.

Now in order for reality or any part of it to become known to us, some sort of condition must obtain whereby it becomes an *object* for us. As it may also be put, something must bring it about that the objects composing reality *appear* to us. But the question is: what makes reality into an object for us? Its being an object for us is not established by its simple existence. And whatever allows reality to be an object for us cannot be merely postulated or taken for granted as a primitive fact – it stands in need of philosophical explanation, if anything does.

Whatever it is that allows reality to become an object for us is naturally and perhaps inevitably conceived as some sort of fundamental connecting relation between reality and ourselves. The question is then what this relation consists in. It cannot consist simply in reality's impressing itself on our minds, for in order for this to result in knowledge of reality, there would have to be something about us which made us appropriately receptive to it: our minds would have to be capable of transforming the impress of reality into a representation of it. That is, we would have to be already immanently related to reality. Nor does it help to reverse the story and conceive our contact with reality as the result of our own activity, since in order for our minds to reach out and read off the features of

reality, we would have to know how to locate and read it – and again this condition could not be fulfilled unless reality were already an immanent object for us.

The prospect of circularity or an infinite regress looms. It seems that, if objects are originally independent from the subject, as the natural conception requires them to be, then any description of the relation which connects us with reality will either presuppose what needs to be explained, or require the postulation of a further, more primitive connecting relation in a series without end. No attempt to break the circle, or block the regress, by appealing to some third term independent of us and reality (e.g. God), or by supposing that our capacity to represent reality somehow follows from our inclusion in it, can succeed. To do so would be to appeal to what Kant calls a ‘preformation system’ (B167), a transcendent state of affairs whereby the agreement of our representations with reality is guaranteed prior to our forming them. But we have reason to accept such a hypothesis only if we have a true representation of the system itself (a true idea of God or our place in reality), to assume which is again to presuppose that, the possibility of which needs to be explained. Hence the futility of invoking a *‘deus ex machina’*, as Kant puts it in the letter to Herz. A dilemma then arises: either it must be admitted that we cannot account for our relation to reality, which makes all assertions regarding the nature of reality and our relation to it dogmatic; or the idea that we stand in a knowledgeable relation to reality must be renounced, which is to embrace skepticism. (A third but scarcely more attractive option is to identify reality with the contents of our own minds, i.e. solipsism.)

The problem is rooted in what we are naturally disposed to think. On the story told by pre-philosophical common sense (what Kant calls ‘our common understanding’), there is first of all a set of objects composing a world, into which the subject is then introduced as a further item; when the subject’s eyes are opened and its cognitive functions are in working order, the world floods in and knowledge of the world results. Common sense itself is unable to say how the presuppositions of this story are fulfilled. The epistemologies of pre-Kantian philosophy provide many different attempts to show how they may be fulfilled, and thereby account for our presumed knowledge of reality, but because they all remain within the terms of the story told by common sense, the result, on Kant’s view, is always the same: they all reduce on examination to the bare, non-explanatory claim that we represent real things because they affect us and because we have an immanent capacity to represent them.

It is easily seen how these remarks apply to rationalism and empiricism. For the rationalist, our representation of the world results from the intellect being struck by the rational order inhering in the world; for the empiricist, it is generated by the array of sensory data that results from the impinging of things on our senses. The form of explanation is however in both cases the same, in so far as both epistemologies take for granted the possibility of reality’s becoming an object for us: they do this at the point at which they assume that we have innate ideas or ideas manifest to the light of our reason, the veracity and harmony of which with real things is assured; or, that we are fitted to form sensible ideas, and concepts from those ideas, in such a way as to map the qualities of real things; or, that the order that we discover in our ideas replicates the order of ideas in God’s mind; and so on. The fundamental objection to making assumptions of this kind is not, for Kant, that it leaves room for skepticism – although that is true – but that it signifies a collapse of philosophical explanation at the crucial point. The only reason for

believing that there is a pre-ordained harmony between reality and our representations, or for accepting any other fundamental epistemological principle intended to guarantee that reality is knowable, is the belief that we represent reality: the principles of rationalist and empiricist epistemology lend no support to this assumption, but merely re-express our natural confidence in it; the assertion that pre-ordained harmony or whatever gives us the capacity to represent reality does not make any philosophical advance over the mere assertion that we are capable of representing reality. The upshot is that our conviction of the reality of the objects of our representation is displayed as groundless: we are left unable to say anything about the status of those objects and why we ascribe any degree of reality to them. Hence the continual vacillations of pre-Critical philosophy between skeptical admission that no philosophical account of reality and our relation to it can be given, and dogmatic assertion regarding the nature of reality and our relation to it. For these reasons, it may be held that pre-Kantian epistemology does not so much attempt to solve as fail to recognise the problem of reality.

The problem of reality does not rest on any special assumptions regarding the necessary conditions of knowledge or the nature of cognition. The argument is not, for example, that human cognition inevitably puts its stamp on objects in such a way as to make it impossible for reality to survive its filtration through our medium of representation. What it presupposes is only that we have the bare distinction between a real thing and an object of representation, and that neither concept implies the other. The essential point is that, just as we lack any positive reason for believing that our representations do not match reality, that reality is *not* open to being represented by us, so we also lack any reason for believing that it *is* open to being represented by us – that reality is what our representations are of. The two assertions are equally groundless. The underlying problem is that, although there is nothing contradictory in the idea of there *being* a fundamental connecting relation that allows reality to become an object for us, in order for us to *represent* this relation, as would be required for philosophical knowledge of our relation to reality, we would need to stand outside our capacity for representation, which we cannot do.

That the assumption of a match between our representations and reality is natural and compelling, and even that there may be something unintelligible about the idea of there not being any such relation, is agreed by Kant. The question is whether we can have any rational insight into this fact, if it is one. What the foregoing suggests is that in order for the assumption of a connection between subject and object to be validated, it is necessary for philosophical reflection to depart from the realist story told by common sense. The fact that common sense is unable to conceive any alternative to the assumption that reality is known to us, other than skepticism, is thus beside the point; what it means is just that the most that we can be asked to do, in advance of seeing what a philosophical alternative to realism would look like, is to suspend our instinctive commitment to realism.

Nor, it will be seen, does Kant ultimately have any quarrel with the realist form of explanation as such. That it is legitimate to refer at some point in the explanation of our knowledge to the fact that things simply are so and so – the pattern of explanation shown most clearly in the case of simple perceptual knowledge – is accepted by Kant: in so far as we remain within the orbit of common sense, it is correct to say that it is because the objects which we perceive really exist, that we have the representations of them that we

do. What Kant rejects is realism at the level of philosophical explanation: the possibility of there being objects for us, things that we can have experiences of and thoughts about, sets a problem which the concept of reality does nothing to help to solve.

Kant's Copernican revolution

This sketch of the problem of reality gives an idea of Kant's motive for reconceiving objects as conforming to our mode of cognition: on Kant's view, to conceive the objects that we cognise as independent from us – the presupposition of 'transcendental realism', in Kant's terminology, to be explained later – is to render the relation of the subject to its objects unintelligible. If, therefore, there is an alternative to realism that can explain how objects are possible for us whilst upholding everything that common sense affirms against the skeptic, there will be every reason for regarding the problem of reality as warranting the abandonment of realism. This alternative, Kant argues, consists in the radical change of methodology which he introduces under the title of a Copernican revolution in philosophy, and constitutes the correct response to the problem of reality.

Pre-Copernican philosophical systems, according to Kant, set out by assuming a domain of objects which are conceived as having being, and a constitution of their own – a class of real things. In this sense, previous philosophical systems are one and all realist. (This generalisation includes, strange though it may sound, idealism of George Berkeley's (1685–1753) sort, because the 'real things' in question may be mental. Hume too is included.) To proceed in this way is to help oneself to the notion of reality, and also to presuppose *ab initio* that we are in possession of a concept of object which has reference independently of the conditions under which we may cognise objects. In Kant's terminology, pre-Copernican, realist philosophy begins by ascribing reference to 'the concept of an object in general'. Having put this concept into play, it then considers how we may take ourselves to stand in relations of knowledge to (at least some) members of the class of real things. Showing this to be so is the task of epistemology, the cost of its failure being skepticism. It is a consequence of this way of proceeding that the concept of an object is fundamentally independent of any epistemological conditions: an object is simply an individual that has being and a constitution, and any epistemic relations that it may have to subjects are to that extent inessential to it. Objects in the same sense are thus in question whether they are known or unknown, knowable or unknowable; whether an object is known or knowable depends upon the experiential history and cognitive capacities of subjects, and has nothing to do, essentially, with what it is to be an object.

The alternative is to begin by making an absolute separation between the supposition that there is such a thing as reality, and the conception of objects which we are capable of cognising. The idea of a thing as it is constituted in itself, a fully real thing, is allowed to stand, but, because it cannot help to solve the Critical problem, denied any role in accounting for the possibility of objects for us. In this way, the concept of an object in general is not pre-assumed to have reference, and a class of real things is not posited at the outset. Instead, philosophical concern focuses on the task of explicating the concept of an *object-for-us*, that is, defining the class of knowable objects. Epistemological conditions, the possibility of being known, are thereby incorporated into the concept of an object in so far as we can suppose it to have reference; that the concept of an object

has reference, and that its object is a possible object for us, Kant will try to show, rest upon one and the same set of conditions. What pre-Copernican philosophy treats as two distinct matters – objecthood and knowability – are thus treated as one.

The distinction of epistemology and metaphysics: the ‘transcendental turn’

It is a consequence of this transformation in the concept of an object that Copernican philosophy revises the relation between metaphysics (or ontology) and epistemology, and in a sense blurs the boundary between them. In pre-Copernican philosophy, there is a clear conceptual division between the question of metaphysics/ontology (what is the constitution of reality?) and the question of epistemology (how do we attain knowledge of reality?). These two sets of concerns are bound to be intermixed in any worthwhile philosophical system, but they remain from the pre-Copernican point of view separable in principle, due to the detachability of knowability from objecthood to which pre-Copernican philosophy is committed. Kant’s transcendental question concerning the possibility of objects – as expressed in the letter to Herz: ‘What is the ground of the relation of that in us which we call “representation” to the object?’ – differs from either of the traditional questions, precisely because the philosophical enquiry to which it leads is intended to undo their distinctness. The traditional metaphysical/ontological question is suspended by Kant – fully real things are not objects that we can intelligibly seek knowledge of – and the sense of the epistemological question revised accordingly. The transcendental question concerning the conditions under which objects are possible for us is therefore not equivalent to a question about the conditions of being, or to a question about the conditions under which objects can be known, and cannot be resolved back into either of them (or their conjunction).

The Copernican revolution is often identified with an ‘epistemological turn’ in philosophy, meaning that it considers all metaphysical questions from an epistemological, justificatory angle (it replaces ‘the question of fact (*quid facti*)’ with ‘the question of right (*quid juris*)’, as Kant puts it (A84–5/B116–17)). This formula points to something important and genuinely present in Kant’s project, but it fails to capture the sense in which it is also intended to change the very framework within which epistemological questions are understood. It also obscures the important point that, because Kant’s transcendental question differs from the traditional question of epistemology, it follows that, for Kant, *even if* epistemology could demonstrate that our cognitive relation to objects is immune to all of the familiar forms of skeptical doubt, it would still not supply what is most fundamentally needed philosophically, because it would still not have dealt with the question of what makes it possible for a real thing to become an object for us. It is, in fact, Kant’s view that pre-Critical epistemologies are debarred from providing the kind of skeptic-proof justification for knowledge claims to which Descartes aspired, and that only transcendental philosophy can rectify this situation; but the motive for transcendental philosophy lies in a demand for philosophical explanation which is independent from Descartes’ quest for certainty. In truth the epistemological turn is only one aspect of Kant’s more wide-reaching *transcendental* turn.

Idealism

It is evident that the subtle but far-reaching adjustment to the concept of an object described above, which lies at the base of the Copernican solution to the Critical problem, implies straightforwardly a rejection of realism. If we now return to the original Copernican claim that objects should be reconceived as conforming to our mode of cognition, the positive commitment of Copernicanism to an idealist conception of objects emerges. To suppose that objects must conform to us is to reverse the customary direction of explanation of knowledge. In the realist scheme, the arrow of explanation runs from the object to the subject: if a subject S knows an object O, then the explanation for S's representing O lies ultimately in O's being the way it is; had O not existed or been otherwise, S would not have represented O or would have represented O differently. Kant reverses the arrow: the deepest, most abstract and encompassing explanation of representation lies in how S is. The constitution of objects is thus determined at the most fundamental level by the subject. And it is a corollary of this pattern of explanation that the subject is *active* in knowing objects. In order for the Copernican claim that objects must be regarded as conforming to our mode of cognition to be made good, the subject must be thought of as *making it the case* that objects conform to its mode of cognition, and this it can do only if it carries over its own constitution to the side of the object, i.e. in some sense actively produces the object. (Otherwise the story will be incomplete: a gap will remain between the subject's having such and such a constitution, and its object's being such as to conform to it.) As Henrich puts it, the principle by which the world is accounted for is 'incomprehensible' unless it 'executes the activities' from which the world originates.

The general approach of Copernican philosophy in answering the question of how objects are possible for us, is therefore to say that, in a recondite philosophical sense, the subject *constitutes* its objects. It maintains, furthermore, that these subject-constituted objects compose the only kind of reality to which we have access: reality in the stronger sense of a realm of objects constituted independently of the subject may be admitted as something that we can (perhaps, must) conceive, but knowledge of it is held to be impossible. On this approach, skepticism is refuted by showing that, although claims to knowledge of real things in the strong sense must, as the skeptic says, be rejected as dogmatic and groundless, reality in the weaker sense is something that we can know precisely because we constitute it. Knowledge claims are thus defended on the basis that reason can have insight into 'that which it produces after a plan of its own' (Bxiii).

By drawing the analogy with Copernicus (Bxvi, Bxxii[n]), Kant does not mean therefore that transcendental philosophy demotes man from a position of centrality in the cosmos, in the way that Copernicus' discovery may have been felt as doing; in fact it has precisely the opposite – humanistic – implication that we stand at the centre of the natural world. Kant means by the comparison that his philosophy, like Copernicus' heliocentrism, explains what appears to be a wholly objective phenomenon in subjective terms: just as Copernicus explains the *apparent* movement of the sun in terms of the movement of the observer on the earth, Kant explains our knowledge of *apparently* independently constituted objects in terms of our mode of cognition. In both a phenomenon which had been regarded previously as having independent reality is

redescribed as an appearance, dependent on the subject. In that respect both Kant and Copernicus break with common sense.

Kant's Copernican strategy immediately raises a question. If the subject constitutes its objects, how much of the object is the subject responsible for? How, indeed, can Kant's Copernicanism avoid collapsing objects in their entirety into the mind that experiences them, as in Berkeley's *esse is percipi*?

Kant has a clear and deep answer to this question, which provides him with a principled reason for claiming that his idealism is *prima facie* distinct from Berkeley's. Since, for Kant, the philosophical motivation for regarding objects as subject-dependent derives from the problem of reality, and not from the kind of considerations that move Berkeley, there is reason for regarding objects as subject-dependent *only* to the extent that they are conceived in terms of the conditions under which objects for us are possible at all, i.e. only with respect to those of their features by virtue of which they conform to the structure of experience; we are justified in regarding as subject-dependent only whatever in objects pertains to the possibility of their being objects for us at all. The writ of idealism runs no further. Crucially, it therefore does not extend to the existence of objects: 'representation in itself does not produce its object in so far as *existence* is concerned' (A92/B125). And in the *Prolegomena* Kant says that his Critical idealism, unlike Berkeley's idealism, is not after all a 'genuine' idealism, because it concerns not the existence of things but only the properties that we predicate of objects by virtue of which we can know them (289, 293–4).

To say that the subject constitutes its objects is therefore not to say that objects are *created* by our representations. The causing of objects by representations is in fact a form of knowledge that can be ascribed only to God. For us, the relation of representation to its object involves a complex mix of passivity and activity, and because our representations are neither simple effects nor simple causes of their objects, it is necessary to explain how object and representation can agree: not being related in the way that causes and effects correspond to one another, their relation needs to be specified, a task that occupies the first half of the *Critique*. The conformity of objects with our knowledge is not therefore assured at a single stroke by the Copernican hypothesis: the problem confronting realism of saying how representations relate to their objects reappears in transcendental philosophy, but with the difference, it is Kant's contention, that it can now be provided with a satisfactory solution.

In the light of all this, Kant's assumption that there are a priori elements in cognition, and the *Critique*'s exclusive concentration on them, is readily intelligible: the a priori element in cognition as a whole is the object-enabling structure of experience, the set of conditions that makes objects possible for us, and the a priori features of objects are those by virtue of which objects conform to that structure. Once this a priori structure is in place, knowledge becomes an a posteriori affair: objects may be regarded as independent from the subject and the realist model of explanation applied, i.e. our representations explained by objects rather than vice versa. Kant thus *accepts* realism (a subject S represents O because of how O is) at the level of common sense. This, it will be seen, is what Kant's conception of what he calls 'empirical reality', by means of which he seeks to harmonise his Copernicanism with the realism of common sense, amounts to: the empirically real features of objects are those which they have over and above (and conditionally upon) their a priori features, and on the basis of which the realist form of

explanation has legitimate application. Philosophical realism may be regarded as confusing these two levels, the philosophical transcendental, and the pre-philosophical empirical: the realist projects the form of explanation which we employ at the pre-philosophical level onto the level of philosophical explanation, without considering what makes empirical reality possible in the first place.

The concept of a Copernican revolution just described is filled out in Kant's doctrine of transcendental idealism. This says, in the briefest summary, that the object of our knowledge is 'to be taken *in a twofold sense*, namely as appearance [*Erscheinung*] and as thing in itself [*Ding an sich selbst*]' (Bxxvii), and that objects are known to us only in the first sense, as appearance (Bxx, Bxxvi). Here there is a transition, from the bare Copernican precept that objects are to be considered as conforming to our mode of cognition, which is a strictly *methodological*, metaphysically neutral claim about the basis on which philosophical enquiry should proceed, to a non-neutral, substantive metaphysical, idealist claim about what the objects of our cognition *are*. The methodological and substantive claims are connected at a very basic level, because in so far as an object is conceived as an appearance, it is conceived as something which can be considered as necessarily (by its nature) conforming to our mode of cognition; in so far as it is conceived as a thing in itself (as 'real *per se*', Bxx), it is conceived as something which cannot be considered as necessarily (by its nature) conforming to our mode of cognition, but rather as something to which our mode of cognition must conform. The Copernican method commits us to the former, transcendental idealist, conception of objects.

It is a further thesis of Kant's – which surfaces, as we saw in the previous chapter, in the Preface (Bxvii-xx), and is defended in the first half of the *Critique* – that our mode of knowledge is sensible, and thus that objects are possible for us only when given in sense experience. From this it follows that appearances, objects that conform to our mode of cognition, are *exclusively objects of sense experience*. The distinction between appearances and things in themselves thus corresponds to the earlier distinction between the two different kinds of metaphysics: the metaphysics of experience has application to and provides us with knowledge of a reality composed of appearances; transcendent metaphysics attempts to gain knowledge of a reality composed of things in themselves.

'Transcendental'

The philosophical method which runs alongside transcendental idealism bears no resemblance to the rationalist's inspection of clear and distinct ideas or application of the principle of sufficient reason, or to the empiricist's anatomy of sense experience (Locke's *Essay* is a mere 'physiology' of human understanding, according to Kant, Aix). It consists in the identification of what Kant calls 'conditions of possibility', or transcendental conditions. These are variously said by Kant to be conditions of 'experience', of 'possible experience', of 'objects of experience', of 'appearances', of 'knowledge of objects', and so on. These conditions must be fulfilled before the subject can be epistemically related to an object. Kant attempts to show that they include the central tenets of common sense metaphysics, such as that there are substances that persist throughout change and that every event has a cause. The arguments that identify these conditions are called by Kant transcendental proofs. Each identifies a different respect in

which objects must conform to our mode of cognition and so legitimates a different component of the metaphysics of experience. A transcendental proof has the peculiarity that it converts a possibility into a necessity: by saying under what conditions experience of objects is possible, transcendental proofs show those conditions to be necessary for us to the extent that we are to have experience of objects at all.

It is by now quite clear why Kant should call his philosophy an idealism, and it is called ‘Critical’ because it is premised on a prior examination – critique – of our cognitive powers. (Pre-Critical philosophy, omitting this task, fails to ground its claims properly, so that even where its conclusions are correct, it asserts them merely dogmatically.) Kant’s characterisation of his Critical idealism as ‘transcendental’ requires further comment. No simple definition can capture the complex sense which this term acquires in the course of the *Critique*, but its core meaning is again bound up with the Copernican revolution. Kant says in the Introduction: ‘I entitle *transcendental* all knowledge which is occupied not so much with objects as with the mode of our knowledge of objects in so far as this mode of knowledge is to be possible a priori’ (A11–12/B25). And in the *Prolegomena*: ‘the word “transcendental” . . . does not signify something passing beyond all experience but something that indeed precedes it a priori, but that is intended simply to make cognition of experience possible’ (373n). Transcendental is thus not to be confused with transcendent, which does precisely mean ‘passing beyond all experience’ (the distinction is drawn explicitly at A295–6/B352–3). Transcendental enquiry is therefore enquiry into the cognitive constitution of the subject to which objects must conform; its product, transcendental knowledge, is at one remove from objects, and concerns only what makes objects, and a priori knowledge of them, possible.

The contrast of analytic and idealist interpretations of Kant should now make more sense. The analytic interpretation seeks to extract from the *Critique* an account of our most basic conceptual presuppositions, with a view to refuting the skeptic directly. The idealist interpretation by contrast regards the *Critique* as attempting to answer a kind of question not formulated in pre-Critical philosophy, and transcendental enquiry into the possibility of objects as subsuming the task of justifying our knowledge claims. Whereas on the analytic interpretation the *Critique* provides new answers to traditional philosophical questions, on the idealist interpretation it reconceives the framework within which philosophical questions are raised and answered. On the idealist view, what is missing from the analytic interpretation is an account of how the structure of experience relates to reality. Reality must share the structure of experience, if our experience is to be experience of reality. The structure of experience is however not a set of logical truths. Nor, being necessary for experience, can it itself be grounded on experience. Merely to appeal to what is contained in our concept of experience in a sense leaves everything open, for what is to say that anything answers to that concept? That the parts of the structure of experience are mutually supporting, and that we are unable to see how our most basic concepts could be discarded, is not enough. For these reasons, on the idealist interpretation, the question of what secures an object for our conceptual scheme as a whole – Kant’s Critical problem, once again – warrants the Copernican conception of the subject as shaping the world.

Proving the doctrine

Kant's description of the Copernican revolution as a 'hypothesis', with which we may 'make trial', may give the impression that the *Critique* is ultimately founded on a naked choice of methodology, and that as a result transcendental idealism is just a proposal about how objects may be conceived, rather than a doctrine about what they really are. If transcendental idealism did rest on a methodological decision alone, then Kant's project would be hollow in an important sense, for there would be nothing to show that considering objects *qua* conforming to our mode of cognition has the significance that Kant claims for it. At most Kant would have shown transcendental idealism to be a coherent alternative to other, realist positions, which would remain in themselves untouched. Furthermore, the skeptic could then point out, with justification, that the fact that we *can* tell a story about objects which, if true, would entitle us to claim knowledge of objects, does not mean that the story *is* true and that we *are* entitled to claim knowledge of objects: that the Copernican hypothesis instructs us to proceed *as if* objects are knowable does not show them to *be* so.

A purely methodological Copernicanism would therefore be in its own way every bit as dogmatic as the metaphysics that Kant rejects. It is consequently important that this is not how Kant's argument actually runs. Kant does not intend to merely assume the truth of transcendental idealism at the outset and trace its consequences. Recognising that something positive must be done to establish his metaphysic, Kant describes it as the 'main purpose' of the *Critique*, not merely to articulate, but to *prove* the doctrine of transcendental idealism (Bxxii). Two attempted proofs are presented: an 'apodictic' proof in the Transcendental Aesthetic and Transcendental Analytic, concerned with space and time and the concepts of the understanding (Bxxii[n]); and an 'indirect' proof in the Antinomy of Pure Reason, according to which the assumption that the objects of knowledge are things in themselves leads unavoidably to contradictions (Bxx).

Kant's success in attempting to prove transcendental idealism is highly disputed. For this reason it is crucial to appreciate the difference made by the problem of reality. If the problem goes as deep as Kant believes, then the injunction to consider objects as things inherently conformable to our mode of cognition is grounded on the impossibility of making cognition intelligible on the basis of the only alternative, viz. considering them as things in themselves; Kant may justly deny that his Copernican methodology is arbitrary. And, with the problem of reality located in the background to the *Critique*, Kant does not have to accept the entire burden of proof in his argument with the philosophical realist: realism ceases to count as the default position. This is the point of Kant's statement in the Preface that the history of metaphysics gives us reason to 'make trial' with Copernicanism: though realism has the strength of the incumbent – its hold on our thinking *de facto* is beyond question – its authority *de jure* need not be accepted in philosophical contexts. So although the *Critique* cannot begin by assuming that the objects of our knowledge are mere appearances, it can proceed on the basis that no appeal is to be made to the independent reality of things in order to explain the possibility of objects for us: realism can be legitimately suspended without being assumed to be false. In this respect Kant's procedure is genuinely experimental: it asks us to begin by bracketing our natural realist convictions and to entertain an unfamiliar hypothesis, the

content of which cannot be properly grasped, let alone evaluated, in advance of its systematic development. In this way Kant accommodates the difficulty, noted earlier, that at the outset we can have no clear notion of what it is to abandon realism without abandoning our natural conviction of the reality of the world, and it is left open to him to show that Copernicanism costs us nothing in terms of this conviction.

This point concerning where the burden of proof lies is, it will be seen, particularly important with regard to the first of the two proofs of transcendental idealism in the *Critique*. It also turns out to make a difference to the understanding and evaluation of Kant's more specific arguments in the *Critique*, concerning concepts like substance and cause. If Kant is regarded as at each point having to defeat a presumption in favour of realism, then his arguments are largely unpersuasive and indeed hard to follow. If on the other hand Kant's arguments are interpreted against the background assumption that there is pressure to account for objects in some non-realist set of terms, they regain force and intelligibility.

In conclusion, it will help to recapitulate, and indicate the interrelations of, the various descriptions that have been given so far of Kant's philosophical project and the motivation of the *Critique*.

The tensions within the Enlightenment that Kant grappled with in his pre-Critical writings led him to question the possibility of metaphysics, and in the Preface Kant describes the *Critique* as having the job of deciding if it is possible for us to have metaphysical knowledge. Bound up with the problem of metaphysics is a complex set of concerns. The failed metaphysics of the *Dissertation*, Kant's last attempt at rationalist realism, had led him to formulate the Critical problem, first stated in the letter to Herz, and his view of this problem changed over the course of the silent decade, as he came to see that the difficulty of relating intellectual representations to things 'as they are' is insuperable, and that no account can be given of our relation to reality conceived as subject-independent (in other words, that the problem of reality has no solution). The new task of transcendental philosophy, which replaces ontology and epistemology as traditionally conceived, is to account for the relation of representation and object independently of realism, i.e. to explain how objects are possible at all for us without assuming their independently constituted reality.

The fate of metaphysics therefore hangs on the solution to the problem of transcendental philosophy. Since in metaphysical speculation reason attempts to grasp objects lying beyond experience, what needs to be determined in order for it to be decided if metaphysics in this (transcendent) sense is possible, is the conditions under which objects in general are possible for us. The self-examination of reason which Kant says is required in response to the conflicts of metaphysics thus coincides with the transcendental task of explaining how objects are possible. Both require an investigation into the nature and scope of human cognition.

The Copernican revolution is Kant's answer to the question of how objects are possible and his solution to the problem of metaphysics. The transcendental enquiry which follows from the Copernican hypothesis tells us, in the form of transcendental proofs, how we must constitute objects in order that experience be possible, and so provides us with a priori knowledge of objects. But it does so only on the condition that the objects in question are identified with appearances rather than things in themselves, i.e. on the condition of transcendental idealism. This yields the first verdict of the

Critique, that metaphysics is possible in the (immanent) sense of the metaphysics of experience. All pre-Critical philosophy, Kant maintains, has assumed that the objects of our knowledge are things in themselves, and he claims to be able to show both that knowledge of things in themselves is impossible, and that it is this assumption which gives rise to the contradictions of metaphysics. This yields the second verdict, that transcendent metaphysics is impossible: 'we can know a priori of things only what we ourselves put into them' (Bxviii).