

Strawson on Kant

Jonathan Bennett

From: *Philosophical Review* 77 (1968), pp. 340–349.

Most of the *Critique of Pure Reason* is prima facie dead, because prima facie dependent upon wholly indefensible theories. The commentator's dominant problem is to display the life below the surface: showing where an argument has an innocent analogue yielding the same conclusion, or a weaker but still untrivial one; testing these mutant arguments and conclusions for consonance with the broad lines of Kant's thought; and so on. There have been many attempts and some piecemeal successes, but what has been needed is mastery of the relevant relationships as they obtain at the levels both of fine detail and of over-all plan. This mastery is now available to us. The dominant problem has been solved, in all essentials, by Strawson.¹

His individual sortings of wheat from chaff are almost all absolutely right and soundly defended. His section, 'The Two Faces of the *Critique*', a synopsis of the relationship of Kant's most insidious theory to the salvageable parts of his thought, shows unexampled fairness, clarity, and secureness of grasp. What is most impressive of all, though, is Strawson's ability to hold small points within the setting of the over all picture, moving from one scale to the other and back again without breathlessness. He has made himself at

home in the Kantian intellectual world, and has learned to move easily and naturally in it, yet familiarity has not dulled the sharpness of his perception of what has to be rejected. In particular, his intimacy with 'transcendental idealism' (I here reluctantly follow Strawson's usage) has engendered understanding, and almost sympathy, but no tolerance.

It is easy to agree that transcendental idealism is unacceptable; but to know how it functions in the *Critique*, how much damage it does, how much of what it does is damage, one needs a criticism of it which is accurate, deep, and comprehensive. We have this too, in Strawson's forty pages on 'The Metaphysics of Transcendental Idealism'. His comparison of transcendental idealism with Locke's thesis about primary qualities (pp. 250–256) is superbly done; and here is an example from another part of the book, attacking the transcendental idealist thesis that time is not part of super-sensible reality but merely the guise in which 'we appear to ourselves':

Do we really so appear to ourselves or only appear to ourselves so to appear to ourselves? It seems that we must either choose the first alternative at once or uselessly delay a uselessly elaborated variant of it.

¹ *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. By P. F. Strawson. (London, Methuen & Co. Ltd.; New York, Barnes & Noble, Inc., 1966.

Then is it a temporal fact. . . that we really so appear to ourselves? To say this would be to go back on our choice; for all that occurs in time belongs on the side of appearances. So it is not a fact about what happens in time that we really appear to ourselves in a temporal guise. I really do appear to myself temporally; but I do not really temporally appear to myself. But now what does 'really do appear' mean? The question is unanswerable; the bounds of intelligibility have been traversed, on any standard [p. 39].

That passage also typifies the deliberately stylish manner of the writing. Strawson has not sacrificed intellectual to aesthetic virtue: as well as being surprisingly economical (the quoted passage defies abbreviation), his prose is a perfect vehicle for those limber, precise movements through the Kantian structure which constitute the book's chief glory.

Strawson inevitably implies many conclusions about how particular Kantian passages should be construed or revised, but defends few of them—presumably because a tangle of textual detail would have harmed the book's clarity of line. The reason is a good one; but it should be said that Strawson's abstract level of argument, his uncondescending prose style, and his expectation of readers who are informed and exegetically competent (having 'read and re-read' the *Critique* and carried away at least 'a sense of great insights') make *The Bounds of Sense* a fairly difficult book.

The complexity of the Dialectic—even of what merits salvage in a reconstruction—conflicts with Strawson's desire for 'a clear, uncluttered and unified interpretation'. In the event, I think, the Dialectic suffers: an interpretation preserving more of Kant's twists and turns would have exposed central issues which Strawson leaves veiled. To defend this judgment I should need a whole book; and I permit myself to express it only by way of explaining why I shall not pursue

Strawson's treatment of the Dialectic here, remarking only that in this respect, too, *The Bounds of Sense* is decisively superior to its predecessors.

Agreeing wholly with Strawson's general approach, and with hosts of minor episodes in his book, I look to the middle level of generality for matter for criticism. Here there are three points at which he seems to be open to objections.

The Transcendental Deduction. What Strawson primarily seeks to salvage from the Transcendental Deduction (TD) is an argument from **(1)** 'A self-conscious being must apply concepts' to **(2)** 'A self-conscious being must apply objectivity-concepts'. There is such an argument in the TD, and Strawson has satisfied me that I have erred in seeing the TD's value as wholly residing in Kant's arguments *for (1)*. Still, these arguments are there, and I submit that Strawson errs in virtually ignoring them by treating **(1)** as a 'premise' (p. 87) but not as a conclusion. Admittedly they are closely linked with a theory which Strawson rightly rejects as an 'aberration' (p. 32), but I believe that something of value survives a breaking of the links. Strawson apparently thinks that **(1)** does not need defense anyway: 'Certainly concepts, recognition, some span of memory would be necessary to a consciousness with any experience at all. . . .' (p. 99). At *our* distance from Hume, we may find that safe enough. Strawson continues: '... and all these would involve one another'. That seems safe, too, but it does raise the question: *how* do concepts, recognition, and memory involve one another? The question is a reasonable one (*pace* Strawson, pp. 111–112), and the attempt to salvage something from Kant's defense of **(1)** would at least open up the area in which the answer must lie.

Strawson sees **(2)** as an understatement: in his view, self-consciousness ordinarily requires experience *and bodily occupation* of an objective realm. Wanting to make clear

exactly what is a reason for what, he substitutes for self-consciousness the weaker notion of (self-) reflexiveness, and argues that this requires experience of an objective realm although it may well not be strong enough to require bodily occupation. For x's experience to have 'a certain character of reflexiveness' is for x to be capable of the thought of his experience *as experience*—the thought that 'this is how things are experienced as being' (p. 107), which Strawson apparently connects with the thought of a sensory history as involving 'a temporally extended point of view on the world' (p. 104). Without saying that self-reflexiveness could occur in the absence of personal self-consciousness, and claiming indeed that the former notion 'really does constitute the essential core' of the latter, Strawson maintains that reflexiveness is nevertheless 'something less than' self-consciousness—that is, less than the capability for the thought that 'This is how things are experienced *by me* as being' (pp. 107–108). The expressions 'point of view *on the world*' and 'how *things* are experienced' are not used to beg the crucial question.

I shall use '(2)' now to name the thesis that a being capable of the thought of its experience as experience must perceive an objective realm. Here is one version of Strawson's argument for it (reflexiveness, though not yet explicitly introduced, is clearly the operative notion):

The trouble with such 'objects of awareness' as those offered by the hypothesis [of experience not pertaining to an objective world] is just that their *esse* is, to all intents and purposes, their *percipi*—i.e. there is no effective ground of distinction between the two—so their *percipi* seems to be nothing but their *esse*. The hypothesis seems to contain no ground of distinction between the supposed experience of awareness and the particular item which the awareness is awareness of [p. 100].

This argument assumes that the notion of 'awareness' somehow involves a strong notion of 'awareness-of—*strong* in the sense that it requires an objective accusative. If we ask why the one should in any way involve the other, I think Strawson will answer at the 'deeper level' to which he immediately proceeds: according to thesis (1), 'there can be no experience at all which does not involve the recognition of particular items as being of such and such a general kind'. This requires that we should be able 'to distinguish a component of recognition, or judgement, which is not simply identical with, or wholly absorbed by, the particular item which is recognized'; and it is the requirement for this component which generates, somehow, a requirement for objectivity concepts. I have to confess that I do not really understand this, because I do not understand 'a component of recognition which is not wholly absorbed by the item which is recognized'. One senses that there is something here of great importance, and of a deeply Kantian kind (see p. 97); but Strawson has not adequately laid it bare. He might have done so—I cannot forbear to add—had he explored Kant's neglected arguments *for* (1).

If the 'deeper level' is meant to supplement the surface argument (as we might call it) by supporting the latter's assumption about 'awareness', then Strawson seems to move from 'awareness' to 'recognition' and back rather too easily. Perhaps I am wrong about how the two levels are interrelated. But I think that the most striking episode in the 'deeper level'—namely, the move with 'component of recognition which is not absorbed' and so forth—must be intended as serially linked in some way, in a single long argument, with the rest of Strawson's argumentative material in favor of (2). If not, then I do not see how the latter material can claim (1) as a premise or the TD as a text.

Let us now return to the surface, concede that ‘awareness’ somehow involves ‘awareness-of’ in a strong sense, and ask what the logical shape of this involvement is supposed to be. Must all awareness be strong awareness-of—that is, awareness of something whose *esse* is not its *percipi*? No. That would yield the false and un-Kantian conclusion that all experience must consist wholly of perceptions of objective states of affairs (see p. 101). The only other possibility I can find is this: a being can properly be said to be ‘aware of his inner states, but only by a kind of borrowing from the strong ‘awareness-of’ which he can have in respect of objective items which he perceives. My thought of my present visual field as something of which I am aware is parasitical on my ability to have the thought that I am aware of something objective. Nested within my thought that I am ‘aware of’ an after-image is the thought that my present state could be—although it is not—an awareness of a patch of paint. (The metaphors ‘borrowing’, ‘parasitical’, and ‘nested’ are mine.)

That version of the argument replaces ‘All awareness is strong awareness-of’ by ‘To have the concept of awareness one must have the concept of strong awareness-of.’ The former went too far, but does the latter go far enough? Must a being actually have objective experience in order to have the thought of it, the thought (perhaps) of what it would be like to have such experience? Strawson thinks so:

It may be conceded that each one of us can perfectly well imagine a stretch of his own experience as being such as the sense-datum theorist describes. . . . But of course it is not enough that, equipped with the conceptual resources we are equipped with, we can form such a picture. What has to be shown is that the picture contains in itself the materials for the conception of itself as experience. What has been shown is that it does not [p. 109].

The whole weight of this is borne by ‘contains in itself the materials’, which echoes an earlier metaphor:

The objectivity-condition. . . provides room, on the one hand, for ‘Thus and so is how things objectively are’ and, on the other, for ‘This is how things are experienced as being’; and it provides room for the second thought because it provides room for the first [p. 107].

Some clarification is needed.

When Strawson says that we ‘can perfectly well imagine’ a purely hallucinatory stretch of experience, is he conceding that for thirty minutes this afternoon I might have and be aware of such a stretch? (If not, I am lost.) If so, then the envisaged situation is this: from two until two-thirty I am aware of my sensory states; I therefore have (we concede) a strong concept of awareness-of, and thus the concept of an object of awareness whose existence is not ‘to all intents and purposes’ the same as my awareness of it. I have this latter concept throughout a period when I have nothing to which to apply it. Must not Strawson claim that I can have that concept at that time only because I have *earlier* been in a position to apply it? If not, then again his argument eludes me. But I think he would rest his argument on that claim: there is evidence in *Individuals*—especially in the thesis that a disembodied mind must be a *ci-devant* person—that Strawson really is prepared to argue genetically, basing philosophical conclusions on principles of the form ‘If P, then Q earlier’; which is certainly what he seems to do here. Perhaps he is entitled to, but the literature which argues that he is not deserves an answer.

The Object-Process Argument. The second Analogy contains an argument involving the distinction between **(1)** successively perceiving different parts of an unchanging object and **(2)** perceiving an objective process. Each involves

a subjective succession, yet the distinction, Kant thinks, must be analyzable in subjective terms. He offers something like this:

(1) I perceive different parts of an **unchanging object** =
The successive parts of my perception **could have**
occurred in some other order.

(2) I perceive an **objective process** =
The successive parts of my perception **could not have**
occurred in any other order.

This analysis is inadequate, and anyway Kant mishandles the analysans in (2) by strengthening the necessity of B's not preceding A to the necessity of its following A and thus the necessity of its occurring. By this slide, he reaches the lemma: in treating my sensory states as perceptions of objective processes I bring them (sc. the sensory states) under 'rules' according to which one follows 'necessarily' upon another; from which he concludes that every objective process is causally determined.

Strawson's main criticism seems at first to be aimed at this last step from the rule-governedness of the perceptions to the rule-governedness of the perceived objective processes. Kant's optimism as to its validity, stemming from his general phenomenalism, doubtless needs support from detailed arguments; but one may be surprised at Strawson's saying that the argument ends with 'a non-sequitur of numbing grossness' (p. 137; cf. p. 28). This outburst (like Strawson's dismissal as 'simple-minded' of my favorite objection to Kant's analysis) is explained by the fact that Strawson, having initially got Kant's analysis right, then proceeds to misrepresent it.

Initially: 'Lack or possession of order-indifference on the part of our perceptions is, [Kant] seems to say, our

criterion. . . of objective succession or co-existence' (p. 134). This is correct, if A-then-B's 'lack of order-indifference' is understood as the 'necessity' of B's not preceding A—which reading Strawson himself implicitly endorses on pages 137 and 138. He agrees, then, that Kant's analysis has this shape:

A-then-B is a perception of an objective process =
Necessarily B does not precede A.

Since this is an analysis, it has the form: Necessarily, P if and only if Q, where the operator expresses *analytic or conceptual* necessity. The analysans (Q) also has the form 'Necessarily. . .', and *this* operator expresses something like *causal* necessity.

Returning now to Strawson's explosion. He says:

The character of the fallacy should be clear. . . It is conceptually necessary, given that what is observed is in fact a change from A to B [plus another qualification], that the observer's perceptions should have the order: perception of A, perception of B—and not the reverse order. But the necessity invoked in the conclusion of the argument is not a conceptual necessity at all; it is the causal necessity of the change occurring, given some antecedent state of affairs. . . A conceptual necessity based on the fact of change is [wrongly] equated with the causal necessity of that very change [p. 138].

But Kant has not only a (conceptual) 'Necessarily. . .' governing the whole analysis but also a (nonconceptual, causal) 'Necessarily. . .' in the analysans. Strawson implicitly reduces Kant's analysis to the conditional (he cannot have a biconditional): 'Necessarily, if A objectively occurs before B [and another condition is fulfilled], then the perception of B does not precede that of A'; and this is simply not the analysis which Kant offers. It belies Strawson's own initial exposition,

and his correct remark that the analysis embodies an alleged 'criterion' for something's being (a perception of) an objective process. Kant could not have thought that the attenuated analysis with which Strawson finally credits him yields such a criterion.

Phenomenal Geometry. Strawson argues that Kant's theory of geometry has not deserved 'so abrupt and total a dismissal' as it receives at the hands of 'the positivist view' (p. 278). He sees Kant's theory as including an account of an intelligible discipline called 'phenomenal geometry', the geometry of 'the looks' of things:

Consider the proposition that not more than one straight line can be drawn between any two points. The natural way to satisfy ourselves of the truth of this axiom of phenomenal geometry is to consider an actual or imagined figure. When we do this, it becomes evident that we cannot, either in imagination or on paper, give ourselves a picture such that we are prepared to say of it both that it shows two distinct straight lines and that it shows both these lines as drawn through the same two points [p, 283].

The fact that we cannot do this serves, it seems, to establish or at least confirm the axiom.

Strawson rightly remarks that Kant was wrong to carry over into physical geometry his thoughts about phenomenal geometry, and that the latter are 'wholly inadequate' to bear the theoretical load Kant lays on them. Yet he goes too far: not in saying that Kant was concerned with something like 'phenomenal geometry' but in his 'modest' claim that 'we can take a reasonably sympathetic view of Kant's doctrine' because geometry 'is a complex thing into which different strands or elements enter in connection with one another, and that the strand which receives most emphasis in [Kant's] theory... is not wholly insignificant' (p. 292). Strawson's

defense of this is vulnerable.

(1) I am unconvinced by Strawson's claims for phenomenal geometry's place in mathematical heuristics and pedagogy (pp. 286–287). He finds it 'plausible' to suppose that 'the curious facility with which phenomenal figure-patterns can be elaborated to exhibit... relations between phenomenal spatial concepts' may be 'what underlies the systematic development of geometry as a mathematical discipline'; but he neither defends this nor explains what 'underlying' relation he has in mind. As for the fact that 'anyone who remembers his schooldays remembers learning geometry 'as a body of self-evident truths about spatial figures': this is true, but with luck it will become false. I am told that mathematics could be taught truthfully.

(2) Strawson says that phenomenal geometry treats of 'just such phenomenal figures as can be presented by physical objects in ordinary [perception]' (p. 282). This we may grant; but he goes on to call it 'the geometry of the spatial appearances of physical things', and that is too strong. At best, phenomenal geometry treats of things' visual appearances as given in single, static visual fields—it is the geometry of what can be seen at a glance. Strawson says that 'we must not think of the visual concepts we are concerned with in too static a fashion' (p. 289), but only to allow, it seems, for the production 'step by step' of the single static picture which, when it is complete, verifies the given axiom. If Strawson is not thus restricting himself to single, static visual fields, then it is no longer an 'axiom' of phenomenal geometry that two straight lines cannot meet twice. He implicitly concedes this: he rejects the suggestion that that 'axiom' is refuted by the look of a suitable pair of railway lines, on the grounds that the observer has 'no single picture which he could properly describe in terms of seeing two lines both as straight and as enclosing a space'

(p. 291) ; and this is true only if 'no single picture' means 'no single, static picture'. Strawson's account of how Kant sought to connect phenomenal with physical geometry is admirable (pp. 66–67, 284–285), but if my argument is right then the actual connection between the two is significantly more tenuous than Strawson suggests.

(3) What goes on within phenomenal geometry? A certain axiom is acceptable because it is 'evident that we cannot' give ourselves a certain kind of picture. If this 'cannot' expresses psychological or physical incapacity, then 'evident' becomes suspect and the axiom's status is left unclear. If the 'cannot' expresses ordinary analytic impossibility, then the axiom should be provable, without resort to visual imaginings. Strawson says, in effect, that what is involved is unordinary analytic impossibility:

The axioms are true solely in virtue of the meanings attached to the expressions they contain, but these meanings are essentially phenomenal, visual meanings, are essentially picturable meanings. Any picture we are prepared to give ourselves of the meaning of 'two straight lines' is different from any picture we

are prepared to give ourselves of the meaning of 'two distinct lines both of which are drawn through the same two points' in a way which we count essential to our having pictured what these expressions mean [p. 283].

On the next page we are told: 'The problem of a necessity which is not the result merely of verbal definitions is solved by the theory. . . which shows how necessity may be secured by a phenomenal exhibition of meanings.' I find Strawson's treatment of 'the character of the necessity' involved in phenomenal geometry so deeply unexplanatory, and the notion of 'picturable meanings' so radically unclear, that I cannot argue about them, but can only beg that he will either amplify or retract.

It seems, then, that no adequate case is made out for the significance of phenomenal geometry either (1) as important to mathematics or (2) as closely linked with physical geometry or (3) as able to stand on its own feet as an interesting and intelligible discipline. Why should we not say that phenomenal geometry is 'wholly insignificant'?