

Epistemology and the Cartesian circle

by

ROBERT CUMMINS

(The Johns Hopkins University)

I

Most recent work in English on the Cartesian circle has tended to defend Descartes by arguing that there is no circle at all, or at least that the mistake is not the obvious and crippling one it appears to be. (See [1], pp. 106—122; [2], pp. 245—246; [5], pp. 149—156; [6], pp. 170—180; [9], pp. 193—195; [10], pp. 123—136; and, less recently, [4], pp. 324—338, and [7], pp. 389—390.) Given this tendency to come to Descartes' aid, it is no wonder that little has been said concerning why and how such a brilliant thinker came to make such an incredible mistake. I find this question pressing because it seems to me that the mistake is incredible: it is right on the surface, and it is crippling. Recent attempts to dispel this appearance, though ingenious and enlightening on many matters, are ultimately unconvincing on the question at issue. In what follows, I will attempt to show that the circularity evident in the *Meditations* is the inevitable outcome of a profound and rather subtle mistake Descartes made (and passed on) concerning the epistemological status of epistemology itself.

II

Before we turn to diagnosis, let us briefly rehearse the problem. In the third *Meditation*, Descartes sets out to establish the existence of God. Descartes' stated motivation for this project is that, until he knows whether there is a God, and if there is, whether He is a deceiver, he can never be certain of anything. In the fourth paragraph of *Meditation III* he says:

But every time that this preconceived opinion of the sovereign power of a God presents itself to my thought, I am constrained to confess that it is easy to Him, if He wishes it, to cause me to err, even in matters in which

I believe myself to have the best evidence. And, on the other hand, always when I direct my attention to things which I believe myself to perceive very clearly, I am so persuaded of their truth that I let myself break out into words such as these: Let who will deceive me, He can never cause me to be nothing while I think that I am, or some day cause it to be true to say that I have never been, it being true now to say that I am, or that two and three make more or less than five, or any such thing in which I see a manifest contradiction. And, certainly, since I have no reason to believe that there is a God who is a deceiver, and as I have not yet satisfied myself that there is a God at all, the reason for doubt which depends on this opinion alone is very slight, and so to speak metaphysical. But in order to be able altogether to remove it, I must inquire whether there is a God as soon as the occasion presents itself; and if I find that there is a God, I must also inquire whether He may be a deceiver; for without a knowledge of these two truths I do not see that I can ever be certain of anything.¹ ([8], pp. 158 f., v. I)

I will call this passage the Circle Passage. The problem it raises is simple and obvious: if *nothing* can be known for certain until the existence of a non-deceiving God is established, any argument advanced to establish such a conclusion must rest on premises not certainly known. Hence, the conclusion will be infected with the very uncertainty it is designed to remove, and will therefore be powerless to remove it. On the other hand, if it is legitimate to argue from premises not certainly known, it is not clear why we need bother with the "slight metaphysical doubt" at all. Of course, if the conclusion could be established by *reductio ad absurdum*, there would be no premises to worry about. But Descartes does not in fact argue for the existence of God in this way. And even if he did, there would still be the *inference* to worry about: Descartes is ready to concede at this point that seeing a manifest contradiction in something is not an airtight guarantee of its denial. Hence, argument by *reductio* will not suffice, and in general, he must

¹ It is perhaps worth remarking in passing that this passage is a *prima facie* embarrassment to those (e.g., Frankfurt) who hold that Descartes was only concerned with doubts which were in some sense reasonable and substantial. Descartes plainly says that he has *no reason to think there is a God who is a deceiver*. According to Descartes, then, the whole inquiry concerning God is required to remove a doubt arising from an hypothesis which he has no reason to believe.

concede that his judgments concerning what validly follows from what are subject to doubt in the same way. This shows just how deep the problem goes: the judgment that the conclusion follows from considerations brought to support it will be subject to a doubt removal of which is supposed to be contingent on establishing the conclusion.²

The problem is so obvious, and so devastating, that it becomes almost impossible to resist the idea that Descartes couldn't have meant what he said, or that he couldn't have been saying what he appears to be saying. But we cannot turn without further ado to reconciliationist interpretation, for there are forces at work in Descartes' thought which make some such problem as this almost inevitable.

A less conspicuous problem raised by the Circle Passage is this: the *cogito* appears as an explicit illustration of a belief subject to the metaphysical doubt which it is the business of the inquiry concerning God to remove. This is a problem because the *cogito* is presented in the second *Meditation* as a belief which is not subject to the doubt raised by the Evil Genius hypothesis. Indeed, in the second *Meditation*, Descartes seems to be claiming as the outstanding virtue of the *cogito* that it is above the sort of question raised in the Circle Passage.³

² The problem of circularity is usually raised in terms of *clarity* and *distinctness*: proof of the existence and non-deceptive nature of God is required to guarantee the truth of what is clear and distinct, but the only grounds for accepting the proof is its clarity and distinctness. This way of raising the problem obscures its seriousness: perhaps there are other grounds for accepting the proof, or perhaps in certain special cases clarity and distinctness provide a guarantee which is independent of the existence and nature of God.

³ The crucial passage, of course, is the third paragraph of the second *Meditation* (p. 150 v. I in [8]). But consider also the following: "But it is very certain that the knowledge of my existence taken in its precise significance does not depend on things whose existence is not yet known to me; consequently it does not depend on those which I can feign in imagination." (Paragraph seven of the second *Meditation*, p. 152 v. I in [8])

Descartes is a bit sloppy here. At this point in the *Meditations* the existence of God is not yet regarded as known, yet Descartes can hardly have meant to say that his knowledge of his own existence did not depend on the existence of God. What he must have meant is rather that his knowledge of his own existence

That Descartes makes this claim in the second *Meditation* is just what we should expect. For the *cogito* is what Descartes discovers by applying the method of doubt, and that method is designed, not to refute the beliefs shown to be doubtful, but to refute claims of the form '*p* is not subject to doubt'. And the *point* of this is to discover whether there are any beliefs not subject to doubt. Since the *cogito* is presented as the fruit of this inquiry, it must be regarded as unimpugned by the Evil Genius hypothesis. This is what distinguishes the *cogito* from simple arithmetical beliefs in the second *Meditation*, although these are lumped together in the Circle Passage.

As I see it, then, the thinking underlying the first two *Meditations* is this. To justify a belief is to infer it from something known. Hence, if there is to be any knowledge at all, some beliefs must be known without justification. Now what a justification *does* is remove grounds for doubt. Thus if we can discover a belief for which there can be no ground for doubt, we will have a belief which does not require justification to qualify as knowledge.⁴ In

did not depend logically on any beliefs he had concerning the existence of other beings or things. (The Haldane and Ross translation is not literal where they have "the knowledge of my existence," but it is less misleading than a more literal rendering, e.g., "the notion of myself," as examination of the previous sentence ("Novi me existere quaero quis sim ego ille quem novi.") makes clear.) Perhaps it is worth noting that this passage is *prima facie* embarrassing to those who wish to hold that Descartes attaches one epistemological status to the *cogito* when it is being affirmed, and another when its affirmation is being reflected upon.

⁴ Here I am in disagreement with Frankfurt and others concerning the sense in which Descartes is concerned in the second *Meditation* to claim that the *cogito* is immune to doubt. I cannot believe that Descartes was seriously concerned to claim that it is psychologically impossible to doubt the *cogito* while actually thinking it. *Contra* Frankfurt, it seems to me that Descartes did think indubitability sufficient for knowledge, for in claiming indubitability for the *cogito*, Descartes was claiming that there could be no reason to doubt it, i.e., that it could not coherently be doubted. Understood in this way, the move from indubitability to knowledge is quite plausible, especially when we recall the point made in footnote and above: a ground for doubt needn't be reasonable, it need only be *prima facie* coherent, and not known to be false. A belief for which there could be no grounds for doubt in this sense is a belief which is logically immune from refutation.

Frankfurt gives the show away when, in criticizing Doney, he switches from

order to discover whether there are such beliefs, we assume as true the most powerful and general ground for doubt we can think of—there is an omnipotent being whose sole aim is to deceive—and we investigate various beliefs to determine whether this assumption impugns them in any way. If a belief is impugned by the assumption, it is thereby shown to require justification (for there is a possible ground for doubt which a justification might remove). If it is not impugned by the assumption, then it is what we are after.

Now Descartes plainly means to argue in the second *Meditation* that the *cogito*, and perhaps certain other beliefs concerning one's own state of mind, are not impugned by the Deceiver hypothesis.⁵ This is what distinguishes the *cogito* from belief in the material world, from simple arithmetical beliefs, and from belief in God. Without such an argument, the *cogito* stands undistinguished from the other beliefs which fail the test, and the search bares no fruit. Descartes plainly did not think the search a failure.

What, then, could have led Descartes to say in the Circle Passage that the *cogito* is impugned by the Deceiver hypothesis after all? I think he might have reasoned as follows. "In the second *Meditation* I attempted to argue that the *cogito* is not subject to doubt. But what is the status of that argument? How am I to know it justifies its conclusion? Couldn't the Deceiver deceive me into thinking that I had shown the *cogito* immune to the Deceiver hypothesis? Apparently nothing will be entirely immune from doubt until the Deceiver hypothesis itself is refuted." If this is on the right track, then Descartes is led to make the troublesome claim in the Circle Passage by reflecting on the status of the

"indubitable" to Doney's phrase, "not subject to doubt." He says of Doney, "Evidently he assumes that if something intuited is not subject to doubt, that is the same as its being known to be true." Attributing this assumption to Descartes is not *evidently* mistaken unless we assume with Frankfurt that "not subject to doubt" means "psychologically indubitable." Cf. [6], p. 150.

⁵ This, at any rate, is certainly what Descartes appears to be doing. There is always room for interpretation, but the burden or proof must surely rest with those who would deny Descartes' intention in the second *Meditation* to argue that the *cogito* is not impugned by the Deceiver hypothesis.

argumentation in the second *Meditation*. I think Descartes' worries about this argumentation are ultimately unjustified, and the project he sets himself at the end of the Circle Passage is therefore unnecessary. I do not mean to claim here that no inquiry concerning God is necessary, but only that a proof of the existence of God is not required for the reasons Descartes sets out in the Circle Passage. For it is my contention that Descartes has misunderstood the point of his own argument in the second *Meditation*.

III

In the *Meditations*, Descartes is concerned above all else with the question of what counts as real knowledge. If we are to be clear ourselves about his project, we must distinguish at least three things which might with equal justice all be called theories of knowledge.

- (A) Formulation of a definition of knowledge, i.e., specification of a set of conditions satisfaction of which is necessary and sufficient for something to be known by someone at a time.
- (B) Delimitation of the possible extent of knowledge, i.e., an enumeration and classification of the sorts of things which could satisfy the conditions given in a type-A theory.
- (C) Determination of which of one's current (and important) beliefs satisfy the conditions laid down by a type-A theory.

Descartes had views about all three of these matters. For instance, his view that a belief for which there can be no ground for doubt requires no justification and is therefore known non-inferentially belongs to a type-A theory. His view that beliefs about the material world, and even simple arithmetical beliefs, are subject to doubt and therefore are not known non-inferentially belongs to a type-B theory. And his view that his belief in his own existence is not subject to doubt and is therefore known by him to be true belongs to a type-C inquiry.

Inquiries of the B or C variety obviously depend on prior articulation of at least some results falling under type-A, for, in an

obvious sense, type-B and type-C inquires are applications of a type-A theory. A consequence of this fact is that type-B and type-C theories cannot coherently pronounce on whether the type-A theory they apply is knowable or known. Let us briefly verify this point. Consider any particular type-A theory. Call it k . It has the following general form:

$$(\forall x) (x \text{ is known (by } s \text{ at } t) \text{ iff } Gx).$$

There is no doubt some sense in which we can legitimately ask whether we know k , but we cannot coherently attempt to settle this question within a type-C theory which employs k as its criterion of knowledge. To attempt this would be to attempt to determine k 's epistemological status by determining whether Gk . But any reason one could have for wanting to determine whether k is shown would equally be a reason for thinking that establishing Gk would not show it: doubts about k render G useless as a criterion. In general, establishing that Gx for some x establishes that x is known only if k is supposed above the sort of question about x which could be settled by showing that Gx . It might be that k is *true* in its full universality, but k cannot provide a universal criterion. If k is true in its full universality, and if Gk , then k is shown. But establishing that k satisfies the condition that k lays down would not *establish* that k is known.

I shall need a name for the mistake made by someone who attempts to establish that he knows a type-A theory by applying it to itself, so I dub the mistake "internalism". The worry Descartes expresses in the Circle Passage concerning the *cogito* derives, I think, from his having fallen prey to internalism. In order to put off temporarily questions of interpretation, I will approach this claim by discussing the reasonings of a fictitious philosopher, Descartes*, whose resemblance to Descartes will be taken up in the next section.

Let us see how Descartes* became a victim of internalism. The method of systematic doubt was initiated by Descartes* to discover whether there is a proposition such that, if it is believed by someone s , then s cannot have a reason to doubt that belief. So the target, give or take a nicety, is (a).

- (a) There is a proposition such that, if it is believed by someone *s*, then *s* cannot have a reason to doubt that belief.
- (a) was of interest to Descartes* because he thought belief in such a proposition would be a case of non-inferential knowledge. So Descartes* also holds (b).
- (b) If *s* has a belief for which *s* cannot have a reason for doubt, then that belief counts as non-inferential knowledge for *s*.

Descartes* attempted to establish (a) by actually exhibiting a proposition—call it *p* for now—and arguing that it satisfies the condition laid down in (a). Having done this, Descartes* was in a position to reason as follows: I believe *p*, and, as my argument for (a) shows, I cannot have a reason to doubt this belief. By (b), therefore, this belief counts as non-inferential knowledge for me. Hence, *p* is true.

This reasoning applies a type-A theory—(a) and (b)—to establish that a certain belief is a case of non-inferential knowledge. But it might easily seem that *p* is being inferentially *justified* on the basis of certain premises including (a) and (b). This was in fact how Descartes* came to regard the matter, and he was thus led to wonder whether the inference really justifies its conclusion. "Since inferences justify their conclusions only when their premises are known," Descartes* reasoned, "the question arises as to whether (a) and (b) are known. Now if they are known, they are either inferred or immune to doubt and hence known non-inferentially. They are by no means immune to doubt: indeed, it was the business of parts one and two of my essay to argue for (a) (by exhibiting *p*), so (a) is inferred. But neither are they inferred from something which is immune to doubt. Belief in *p* cannot yet be pronounced non-inferential knowledge because this conclusion depends on (a) and (b) which are not themselves above suspicion. Perhaps I am deceived in accepting them, or in thinking I have reasoned from them correctly."

By now it is obvious that Descartes* has become entangled in internalism. He establishes (a) by exhibiting *p* and arguing that *p* satisfies the condition laid down by (a). This allows him to conclude

he knows *p* *via* (b), for he has just shown that belief in *p* is immune to doubt. Reflecting on this whole process, Descartes* ironically misconstrues it as an attempted inferential justification of his belief in *p*. Construed in this way, it is found wanting because its premises (including (a) and (b)) are not immune to doubt, being impugned by the Deceiver hypothesis. Having come this far, Descartes* concluded that nothing can be immune to doubt and hence known non-inferentially until the Deceiver hypothesis is disarmed, and he devoted the third part of his essay to this task. Scholars subsequently pointed out that Descartes* is involved in a circle: the Deceiver hypothesis is bound to undermine any argument designed to disarm it.

Given this misconstrual of the argument, Descartes* is committed to internalism. For one he treats (a) and (b) as premises in a justificatory argument for *p*, he *must* ask whether (a) and (b) are immune to doubt or inferred from something which is immune to doubt. This is internalism: Descartes* is trying to determine the epistemological status of his type-A theory by applying it to itself. Once committed to internalism, Descartes* is bound to get involved in a circle somewhere, internalism being the inherently circular attempt to guarantee a criterion of knowledge by showing that it satisfies itself.⁶

IV

How faithful is Descartes* to Descartes? This is a difficult question. My story about Descartes* embodies a certain hypothesis about why Descartes wrote what he did. This hypothesis has several things to recommend it. First, it has the advantage of explaining how Descartes came to write the Circle Passage without assuming that he didn't mean it or that it doesn't say what it seems to say. This is a considerable advantage: the *Meditations* present "mature" ideas, ideas which Descartes had worked on for years before he wrote the *Meditations*. We have every reason to believe that their expression in the *Meditations* was carefully

⁶ One might be an internalist, notice the circularity, and endorse skepticism as a consequence. This, though mistaken, wouldn't actually involve arguing in a circle.

considered. Second, my hypothesis explains how what seems to be a glaring error could have seemed inevitable and right to Descartes. Finally, my hypothesis explains these things by attributing to Descartes a mistake—internalism—which would have been extremely natural for him given the way the *Meditations* develops. The first point must stand on its merits, and the second point has already been discussed, so I will turn briefly now to the final point.

Having employed the method of doubt to discover a belief immune to doubt, Descartes naturally turned to the question of what could be justified on the basis of this belief. What makes this natural is just that the search after non-inferential knowledge was undertaken precisely because such knowledge is required as the basis of inferential knowledge. According to my account, the trouble starts when Descartes misconstrues an argument designed to establish that a certain belief satisfies the criterion for non-inferential knowledge as an inferential justification of that belief, i.e., when he puts the argumentation leading up to the *cogito* on a par with arguments designed to justify “derived” beliefs. The mistake is a fairly natural one: when one asks whether a certain belief counts as knowledge, one may either be asking for an argument justifying the belief, or for an argument establishing that the belief satisfies a type-A criterion. In both cases, the belief is shown to be a case of knowledge only if the argumentation meets certain standards, and since the conclusion is the same in both cases—such-and-such is known—it is not surprising to find someone attempting to apply the *same* standards in both cases.

For Descartes, such a mistake would have been especially natural for several reasons. (i) The method of doubt can be seen either as a method of discovery—a way of finding candidates for non-inferential—or as a method for criticizing alleged justificatory arguments. Now all argumentation is supposed to justify its conclusion in some sense, and it is rather natural to apply the critical method to the argumentation involved in using the method of discovery. Since Descartes *argues* for the conclusion that the *cogito* is known, it is natural for Descartes to see this procedure as *justifying* the *cogito*, and hence to require that the argumentation withstand the same critical scrutiny that justifications generally

must withstand. In practice, this means for Descartes that the premises of the argument must be immune to the doubts raised by the Deceiver hypothesis. Once this inquiry is off and running, however, internalism is unavoidable, for one of the "premises" in question is just that beliefs immune to doubt count as knowledge without inferential justification.

(ii) Descartes' working criterion of knowledge—something is known if it is immune to doubt or inferred from something which is immune to doubt—is not the same as his "official" criterion—clarity and distinctness. He never gets around to discussing the problem of a criterion until well after the damage is done, and hence never has explicit occasion to distinguish his type-A remarks from his type-C remarks until well after explicit reflection on the distinction could have helped him avoid the problem. Since Descartes' official procedure is to attempt to extract a criterion from previous discussion of what he knows, and since the criterion he extracts is not the one which is actually functioning in that discussion, it is not surprising that he is insensitive to the role actually played by immunity to doubt in the discussion of the first two *Meditations*.

(iii) There is a genuine problem about the status of a type-A theory operative in establishing type-C claims, and hence a genuine problem concerning the argumentation of *Meditation* two. It is part of Descartes' genius that he realized that this argumentation required some sort of defense. A lesser mind would have avoided the trouble by missing the problem which starts it off.⁷ But in asking whether this argumentation rested ultimately on something immune to doubt, Descartes seriously mistook the sort of defense required. Once the question is asked in this way, a circle is imminent. The question cannot be answered when it arises in this way. It must be criticized and replaced by a similar looking but entirely different question: what is the (or a) correct type-A theory, and how are such theories to be defended? I don't know how this

⁷ This helps somewhat to explain Descartes' resistance to the circularity charge: he feels he is responding in the only way possible to a genuine problem. Descartes, I think, merely *sensed* this problem, without ever fully articulating it and hence without being able to explain clearly to his critics what he was up to.

problem *should* be approached but it is clear that it should not be approached in the way Descartes approached it.

One further point should be made in connection with the hypothesis offered here. In attributing (b) to Descartes, it might seem that I have interpreted Descartes in a way which is shown to be mistaken by the following passage in the *Replies to Objections II*.

To begin with, directly we think that we rightly perceive something, we spontaneously persuade ourselves that it is true. Further, if this conviction is so strong that we have no reason to doubt concerning that of the truth which we have persuaded ourselves, there is nothing more to inquire about: we have here all the certainty that can reasonably be desired. What is it to us, though perchance someone feigns that that, of the truth of which we are so firmly persuaded, appears false to God or to an Angel, and hence is absolutely speaking, false: What heed do we pay to that absolute falsity, when we by no means believe that it exists or even suspect its existence? We have assumed a conviction so strong that nothing can remove it, and this persuasion is clearly the same as perfect certitude. ([8], p. 41, v. II)

Here Descartes seems to concede that a belief for which someone cannot have a reason to doubt, though justifiably believed, may yet be false. And this concession is incompatible with (b). (Cf. [6], p. 179.)

I'm not sure how to understand this passage. Perhaps Descartes really was prepared from the start to allow that a belief for which one cannot have a reason for doubt may yet be false. Or perhaps he adopted this view as an afterthought under pressure of criticism. However, this may be, the interpretation I have offered is easily made compatible with the passage just cited. All that is required is substitution of 'justifiably believed' for 'known' throughout. The problem raised by internalism arises in an exactly analogous way: there cannot be a perfectly universal criterion of the justifiably believed since establishing that the criterion satisfied itself would not establish that the criterion is justifiably believed. My account of the thinking underlying the first two *Meditations* survives the change as well: the problem is to defeat the skeptical view that all belief is arbitrary since it is inferred ultimately from unproved premises by establishing that there is a

proposition—the *cogito*—which is justifiably believed without justificatory argument (i.e., non-inferentially). And the mistake is the same: Descartes mistakes his argument for this conclusion for an inferential justification of the *cogito*. So the interpretation offered here is in no way dependent on the view that, for Descartes, what is immune to doubt is true. I have couched the discussion in terms of what is known (hence true) rather than in terms of what is justifiably believed because I think that this is, in fact, more faithful to Descartes' considered views. But that is another story.

References

- [1] ALEXANDER, ROBERT, "The problem of metaphysical doubt and its removal," in [3].
- [2] BECK, L. J., *The metaphysics of Descartes* (Oxford: Oxford University Press, 1965).
- [3] BUTLER, R. J. (ed.), *Cartesian studies* (New York: Barnes and Noble, 1972).
- [4] DONEY, WILLIS, "The Cartesian Circle," *The journal of the history of ideas*, vol. 16 (1955).
- [5] FRANKFURT, HARRY, "Descartes' validation of reason," *American philosophical quarterly*, vol. 2 (1965).
- [6] FRANKFURT, HARRY, *Demons, dreamers and madmen* (New York: Bobbs Merrill, 1970).
- [7] GEWIRTH, ALAN, "The Cartesian Circle," *The philosophical review*, vol. 50 (1941).
- [8] HALDANE, ELIZABETH, and G. ROSS, *The philosophical works of Descartes* (New York: Dover, 1955).
- [9] KENNY, A., *Descartes* (New York: Random House, 1968).
- [10] TWEYMAN, STANLEY, "The reliability of reason," in [3].

Received on September 27, 1975.