

The Fourth Meditation and Cartesian Circles

C.P. Ragland
Saint Louis University

Everett Fulmer
Loyola University New Orleans

FORTHCOMING
PLEASE CITE OFFICIAL VERSION

Abstract

We offer a novel interpretation of the argumentative role that Meditation IV plays within the whole of the *Meditations*. This new interpretation clarifies several otherwise head-scratching claims that Descartes makes about Meditation IV, and it fully exonerates the Fourth Meditation from either raising or exacerbating Descartes' circularity problems.

Introduction

In the Synopsis of the *Meditations*, Descartes says: "In the Fourth Meditation it is proved that everything that we clearly and distinctly perceive is true" (AT 7:15 / CSM 2:11).¹ Descartes initially proposes this "Truth Rule"² in the Third Meditation, but immediately doubts it: for all he knows, he may have been created by a deceiving God to have *false* "clear and distinct" perceptions. Descartes then seeks to remove this doubt with an argument stretching across the Third and Fourth Meditations. First, he shows that various clearly and distinctly perceived truths entail the existence of an

¹ References to Descartes employ the following abbreviations:

AT: Charles Adam and Paul Tannery, eds., *Oeuvres de Descartes*, 2nd ed., 11 vols. (Paris: Vrin/C.N.R.S., 1974-86).

CSM 1: *The Philosophical Writings of Descartes, Volume I*, trans. Cottingham, Stoothoff, and Murdoch (Cambridge: Cambridge University Press, 1985).

CSM 2: *The Philosophical Writings of Descartes, Volume II*, trans. Cottingham et al (Cambridge: Cambridge University Press, 1984).

CSMK: *The Philosophical Writings of Descartes, Volume III: The Correspondence*, trans. Cottingham, Stoothoff, Murdoch, and Kenny (Cambridge: Cambridge University Press, 1991).

Translations are from CSM or CSMK unless otherwise noted.

² Descartes states this "rule for establishing truth," (AT 7:70 / CSM 2:49) many times in the *Meditations* and elsewhere, usually with some variation in wording. For example: "whatever I perceive very clearly and distinctly is true" (AT 7:35 / CSM 2:24), "everything of which I am clearly aware is true" (AT 7:65 / CSM 2:45), "I am incapable of error in those cases where my understanding is transparently clear" (AT 7:70 / CSM 2:48), and "we . . . avoid error when we confine our assertions or denials to what we clearly and distinctly perceive should be asserted or denied" (AT 8a:17-18 / CSM 1:204). See also AT 7:70 / CSM 2:48 quoted below in footnote three, and AT 7:62 / CSM 2:43, discussed at length later in the body of this essay.

omnipotent God (Third Meditation). Then, he shows that God guarantees the truth of clear and distinct perceptions (Fourth Meditation).

Descartes' doubt of, and argument for, the Truth Rule raises the famous problem of the Cartesian Circle: it seems that he cannot trust his clear and distinct perceptions until he has proven the Truth Rule, but he cannot prove the Truth Rule without first trusting his clear and distinct perceptions. Perhaps because Arnauld originally formulated the Circle as a problem about being sure of God's existence,³ traditional commentary has tended to focus on the Third Meditation's role in the problem. But the Fourth Meditation, too, plays a role that deserves scrutiny.

We consider here two different ways in which the Fourth Meditation may raise circularity problems for Descartes. First, merely by arguing for the Truth Rule, the Fourth Meditation may rule out non-circular interpretations of Descartes' overall project. Such interpretations limit Descartes' aim to the removal of doubts about clear and distinct perception, but a full-blown proof of the Truth Rule seems to reach for more than a mere removal of doubt.⁴ Second, Descartes' *way* of arguing from God to the truth rule may be fatally flawed. On close examination, the Fourth Meditation proof seems to rely on a premise that is merely a strengthened version of the Truth Rule. But if so, then Descartes' argument uses the Truth Rule to prove itself. In contrast to the traditional "external" circle sketched above, this would be a much more obviously vicious "internal" circle.⁵

Addressing these problems, we argue that the Fourth Meditation does not *add* to Descartes' circularity issues. The proof of the Truth Rule is compatible with non-circular readings of Descartes' project. And the details of Descartes' proof do not really involve him in ("internal") premise circularity. This is not to say that the *Meditations* as a whole avoids vicious circularity. Perhaps the Third Meditation or other texts still sufficiently motivate a circular reading of Descartes' overall project. But if so, the Fourth Meditation adds nothing to that motivation. Despite initial appearances to the contrary, the Fourth Meditation, at least, is perfectly innocuous.

I. An External Circle?

³ "Before we can be sure that God exists, we ought to be able to be sure that whatever we perceive clearly and evidently is true." Fourth set of Objections (CSM II: 150/AT 7:214).

⁴ This worry is clearly articulated in Hatfield (2014: 208-09).

⁵ In the external circle, Descartes could not be *justified in believing* the premises from which he derives the truth rule unless he were already justified in believing the truth rule itself, though these premises are all distinct from the truth rule. In the internal circle, the conclusion is itself among the premises. This internal/external terminology follows Carriero (2009: 463, note 40) and Della Rocca (2011: 98). Della Rocca (2006, 2011) raises the worry about internal circularity.

Whether Descartes is guilty of the traditional external circle depends on the nature of his overall project—on what he was trying to show about the faculty of human reason. Gary Hatfield (2014) helpfully distinguishes four different interpretations of Descartes' ultimate aims.

1. **“Certainty, not truth”**: Clear and distinct perceptions make us psychologically certain. While we are having them, we cannot doubt them. Descartes' goal is merely to extend psychological certainty to claims that are no longer clearly perceived (181). He is not trying to prove the truth of anything.

2. **“Remove the doubt”**: Similarly, rather than trying to show that clear and distinct perceptions are true, Descartes merely tries to show that there are no reasonable grounds for doubting them (182-83). His goal is simply to *defend* reason against skeptical assaults.

3. **“Presumption in favor of the intellect”**: Descartes presumes that “we should trust the best use of our cognitive faculties to yield truth, barring some compelling reason not to . . . Removing the grounds for doubt then banishes such reasons, and we are left with our presumption” (183). Descartes' goal is to prove the truth of clear perceptions, but he can do this merely by removing doubts, since the burden of proof is “squarely on the doubter” (183).

4. **“Strong validation”**: Descartes seeks to prove “that reason is sufficient for establishing metaphysical truths about the way things really are,” but he assumes that “the burden of proof remains on reason, not the skeptic” (99). He aims to overcome this burden—to move from distrust to trust—by not only removing doubts, but also giving an additional positive argument on behalf of reason's reliability.

Though the first three readings each face interpretative and philosophical difficulties, the particular problem of circularity afflicts only the strong validation reading. Hatfield nicely expresses the circularity of strong validation as follows:

If reason's trustworthiness has been successfully challenged . . . how could it ever recover? Since it would stand alone in its own defense, the burden of proof would be insurmountable . . . To achieve strong validation, reason must *prove* its own reliability

as a source of knowledge about the basic structure of reality. How could it, when that reliability is the very matter in question?” (99).

On the other three views, however, Descartes’ argument for the truth rule does not move out beyond reason to reality; it stops once Descartes has achieved the more modest goal of removing doubts (on the “presumption” reading, Descartes continues further, but it is the presumption—not his argument—that bridges the gap between reason and reality). His argument aims to show not that the best deliverances of reason correspond to reality, but that they are internally coherent and that doubt about them is unreasonable. Given this more modest aim, there is nothing viciously circular about Descartes’ procedure.

Hatfield contends that “the divine guarantee at the end of the Fourth Meditation apparently runs counter to the first three” of these readings and “supports ascribing a strong validation strategy to Descartes,” because Descartes’ “appeal to God’s certification of clear and distinct perception does more than simply ‘remove the doubt’ that a deceiver might be at work” (208). If Hatfield’s contention is correct, then the Fourth Meditation provides significant—perhaps the *most* significant—evidence that Descartes is guilty of the traditional Cartesian Circle.

Hatfield reads the Fourth Meditation this way because he thinks that Descartes has already removed the doubt in this passage near the end of the Third Meditation:

. . . it would be impossible for me to exist with the kind of nature I have . . . were it not the case that God really existed. By ‘God’ I mean the very being the idea of whom is within me . . . who is subject to no defects whatsoever. It is clear enough from this that he cannot be a deceiver, since it is manifest by the natural light that all fraud and deception depend on some defect.

Descartes first rules out the possibility that his nature might have been produced “by fate or chance or a continuous chain of events” (AT 7:21/CSM 2:14).⁶ Likewise, the idea of a deceiving God suffers from “an internal contradiction: God, who is conceived as perfect, could not be a deceiver” (182). Commenting on this passage, Hatfield says: “If God is no deceiver, then the ‘slight’ and ‘metaphysical’

⁶ This was the first of his principle reasons for doubt in the First Meditation; the other was the possibility that he was created by a deceiving God.

ground for doubt . . . has been removed.” The existence and veracity of God “releases the meditator from the doubt” (177). But if Descartes has removed these reasons for doubt by the end of the Third Meditation, then why does he argue from God to the Truth Rule in the Fourth? It is hard to see why he would do so if he were not pursuing strong validation. The argument for the Truth Rule, Hatfield concludes, makes it “difficult to avoid strong validation” (209).

However, *pace* Hatfield, the Third Meditation alone is not sufficient to completely remove the Deceiving God doubt. For already in the First Meditation, Descartes had considered—and found wanting—an argument from God’s nature to God’s non-deceptiveness:

But perhaps God would not have allowed me to be deceived in this way, since he is said to be supremely good. But if it were inconsistent with his goodness to have created me such that I am deceived all the time, it would seem *equally foreign to his goodness to allow me to be deceived even occasionally*; yet this last assertion cannot be made. (AT 7:21/CSM 2:14)

Since Descartes knows that he sometimes errs⁷, the italicized thought gives him a reason to doubt that God is no deceiver. He cannot completely remove the doubt until he examines this thought more closely and finds it wanting—and that is exactly what he does in the Fourth Meditation. No wonder, then, that in the Synopsis, Descartes says that the results of the Fourth Meditation “need to be known . . . in order *to confirm what has gone before*” (AT 7:15/CSM 2:11; our emphasis).

Because Descartes needs to clarify and defend the Third Meditation’s inference to God’s non-deceptiveness, he does not finally remove the last shreds of doubt until he proves the truth rule near the end of the Fourth Meditation. Two considerations favor this sort of reading.

First, as a matter of logic, demonstrating the incoherence of the deceiving God hypothesis *requires* arguing from God’s nature to the truth rule. Immediately after proposing the truth rule, Descartes says:

. . . whenever my preconceived belief in the supreme power of God comes to mind, I cannot but admit that it would be easy for him, if he so desired, to bring it about that I go wrong even in those matters which I think I see utterly clearly with my mind’s eye

⁷ Consider the opening line of the First Meditation: “Some years ago I was struck by the large number of falsehoods that I had accepted as true in my childhood . . .” (AT 7:17/CSM 2:12).

. . . But in order to remove . . . this slight reason for doubt . . . I must examine whether there is a God, and, if there is, whether he can be a deceiver (*AT 7:36 / CSM 2:25*).

This passage specifies the meaning of “deceiving God”: a creator who brings it about that the Truth Rule is false. Given this meaning, a deceiving God is possible if and only if God’s existence is consistent with the denial of the Truth Rule. But in any classical logic, God’s existence is consistent with the denial of the Truth Rule, if and only if God’s existence *does not* entail the Truth Rule. And conversely, God’s existence entails the Truth Rule, if and only if a deceiving God is impossible. Hence, removing the deceiving God doubt *requires* proving the Truth Rule.

Second, the language at the end of the Third Meditation is somewhat vague and schematic, suggesting a need for further explication. Descartes says that God “cannot be a deceiver, since . . . all fraud and deception depend on *some* defect”: but which defect(s), exactly? Descartes’ extremely terse presentation of the argument does not make explicit its relation to the Truth Rule or to the existence of error. Seemingly aware of the questions he has left open, Descartes immediately declares that after pausing to contemplate God, he will begin “examining *this point* more carefully [hoc diligentius examinem]”. In context, “this point” must be the inference (just stated) from divine perfection to divine veracity. The Fourth Meditation explicates “this point” as an explicit inference from God’s perfection to the Truth Rule that not only specifies the imperfections involved in deception (malice or weakness) (*AT 7:53/CSM 2:37*) but also shows why the inference is not thrown into question by human error.

In short, the Fourth Meditation argument for the Truth rule develops the Third Meditation’s incomplete attempt to remove doubt. It is *not* a further argument that begins after the removal of doubt, but rather is the final step in that removal. It does *not*, therefore, indicate that Descartes was attempting a strong validation of reason.

The strong validation reading gets much of its plausibility from the following train of thought: removing doubt happens inside the mind, while the Fourth Meditation argument aims to prove a correspondence between mind and world (C&D ideas are objectively true); therefore, the Fourth Meditation must go beyond removing doubt. But despite its intuitive pull, the above reasoning is a non-sequitur: proving the mind-world connection is not sufficient to reach beyond the removal of doubt, for two reasons.

First, as noted above, proving the mind-world connection is not a wholly distinct task from removing doubt. The worry that I may “go wrong even in those matters which I think I see utterly

clearly with my mind's eye" is a fear that subjectively certain propositions might fail to correspond to objective truth. Descartes could not *think* this doubt without using the *idea* of the mind-independent world, and so he cannot remove it without similarly using that idea. Second, proving the mind-world connection need not be a more extensive task than removing doubt. The question of whether Descartes pursues strong validation comes down to his intent, not just to the arguments he uses. Proofs can be offered for many different purposes. For example, an attorney may offer the same argument either to prove the truth, or merely to show the jury what they should think, given the evidence. Similarly, Descartes could be using his argument to establish metaphysical truth, but he could also be using his argument merely to show the doubter what she should think, given the best rational assessment of the evidence: she should relinquish her doubts! The text of the Fourth Meditation's proof itself underdetermines which of these (or others) is Descartes' aim.⁸

To sum up, Descartes must have intended the Fourth Meditation's Truth-Rule proof to remove doubt, and may well have intended it to do nothing more. The Fourth Meditation argument, then, does not *by its very existence* raise worries about circularity. However, its means of execution may be a very different story.

II. An Internal Circle?

Let us consider, then, the details of Descartes' argument from God to the truth rule. We begin in section (A) by presenting the argument in its context. Then in section (B) we explain why it can appear circular. Section (C) considers, and rejects, one way of avoiding this circularity. Finally, in sections (D) and (E) we present two reasons for thinking that there is no internal circle.

II.A. The Fourth Meditation's Argument, Problem, Tasks and Solution

Near the opening of the Fourth Meditation, Descartes lays out the opening steps of his argument:

- 1) *Divine Veracity*: "The will to deceive . . . cannot apply to God"(AT 7:53 / CSM 2:37).⁹

⁸ Indeed, Hatfield's four different interpretations of the *Meditations* project are probably best understood as competing theories about Descartes' *intent*.

⁹ Limited as it is to divine intent, Divine Veracity does not by itself imply that the C&D perception is reliable; hence it is not sufficient by itself to remove doubt.

2) *Divine Origin*: I have “a faculty of judgment which ...I certainly received from God”

(AT 7:53-54 / CSM 2:37).

Therefore,¹⁰

3) *Divine Guarantee*: God “did not give me the kind of faculty which would ever enable me to go wrong while using it correctly”(AT 7:54 / CSM 2:37-38).

We cannot err when we judge correctly—that is, in a way that conforms to God’s intentions for the functioning of our faculty.

Immediately after stating this argument, Descartes raises a worry: “There would be no further doubt on this issue,” he says, “were it not that what I have just said appears to imply that I am incapable of ever going wrong. For if everything that is in me comes from God, and he did not endow me with a faculty for making mistakes, it appears that I can never go wrong” (AT 7:54 / CSM 2:38). That is to say, Divine Veracity and Divine Origin appear to entail not only the Divine Guarantee, but also:

Infallibility: “I am incapable of ever going wrong.”

But, as Descartes says: “I know by experience that I am prone to countless errors” (AT 7:54 / CSM 2:38). If Infallibility really follows from Divine Veracity and Divine Origin, then one of them must be false as well. This reason for doubt, raised briefly in the First Meditation, is here expressed more fully.

Descartes’ theological premises face a version of the logical problem of evil.¹¹ To overcome this problem and prove the truth rule, Descartes must carry out two tasks.

First, he must *defend* the truth of his theological assumptions by showing that they do not *really* entail infallibility. This is equivalent to showing that {Divine Veracity, Divine Origin & error} is a

¹⁰ The validity of this inference requires two suppressed premises: one defining a deceptive faculty in relation to correct use, and the other accusing God of deception, if he were to give me a deceptive faculty. We suggest that Descartes was probably assuming something like the following:

Definition of a Deceptive Faculty: If I correctly use a faculty of mine in the production of belief, then if that belief is false, my faculty is deceptive (i.e. is a faculty *for* forming false beliefs).

Transference of Deception to God: If I have a God-given faculty that is deceptive, then God deceives me. Descartes comes close to explicitly stating this latter premise at AT 7:144 / CSM 2:103.

¹¹ A merely probabilistic connection could not cause Descartes to doubt the theological claims, because he takes himself to have a demonstrative proof for them. Suppose that (in raising the objection) Descartes were thinking that error is *very unlikely* (but still possible) given his creation by a veracious God. In that case, when confronted with the fact of error, he would not doubt his creation by a veracious God—for which he has a demonstration—but would simply conclude that something very unlikely had in fact occurred. So the problem is logical, not evidential. Newman (1999, 572-73) and Ragland (2016, Chapter 2) also argue that the Fourth Meditation’s version of the problem of evil is logical and not evidential.

consistent set. To do that, Descartes merely needs to produce one *logically possible* truth-value assignment that would make all statements of that set true. He does not need to *believe* that truth-value assignment. He can execute his defensive task in a purely hypothetical mode, without committing to the claims in his defense.¹²

Second, Descartes needs to *derive* the Truth-Rule from the divine guarantee. For notice: the divine guarantee is not the truth rule, nor does it imply the truth rule all by itself.¹³ To do that, it must be supplemented with a premise linking clear and distinct (C&D) perception to correct use. We call this supplementary premise “Correctness,” and take Descartes to derive the Truth Rule in the following way (continuing from the argument above):

3) *Divine Guarantee*: God guarantees that when judging correctly, I reach the truth.¹⁴

4) *Correctness*: When I assent to what I C&D perceive, I judge correctly.

Therefore,

5) *Truth Rule*: When I assent to what I C&D perceive, I reach the truth.¹⁵

In contrast to the defensive task, this derivative task is not hypothetical. If Descartes is to believe the Truth Rule in the end, he must *commit* to the truth of correctness and the divine guarantee.¹⁶

As the Fourth Meditation continues, Descartes explains the origin of human error in a way that both blocks the inference to Infallibility and establishes the truth of Correctness.¹⁷ He begins by noting that our “faculty of judgment” is really *two* faculties interacting. The intellect displays

¹²Fulmer and Ragland (2017) contains a more substantive examination and defense of this claim.

¹³ Here we agree with Newman (1999: 579-80), but not for the reasons he mentions.

¹⁴ This is a simpler, and equivalent, paraphrase of the direct quote of premise (3) referenced above.

¹⁵ In addition to the passages discussed below (in section II.C), *Principles* I.43-44 clearly gestures at this sort of argument.

¹⁶ Such commitment is compatible with any of the four readings of Descartes’ project discussed above. Even on the “certainty, not truth” reading, Descartes *believes*—affirms as true—all the propositions from which doubt has been sufficiently lifted. His defensive project (of showing that his theological assumptions do not entail infallibility), by contrast, does not require him to affirm any claims.

¹⁷ So his account completes both the defensive and explanatory tasks. Descartes appears to gesture at both these tasks in the synopsis of the *Meditations*: “In the Fourth Meditation it is proved that everything that we clearly and distinctly perceive is true, and I also explain what the nature of falsity consists in. These results need to be known both in order to confirm what has gone before and also to make intelligible what is to come later” (AT 7:15 / CSM 2:11). The defensive task is to remove doubts about the Third Meditation’s grounds for the truth rule, thereby confirming “what has gone before.” His explanatory task “makes intelligible” what comes later in the *Meditations* by clarifying the standard for correct use of judgment (to which he will appeal when discussing the correct scientific conclusions to draw from our sense experiences).

propositions for consideration, and the will passes judgment on them (AT 7:56 / CSM 2:39). The perceptions of the intellect are either perfectly clear and distinct, or involve some degree of obscurity or confusion.

When they are fully clear and distinct, intellectual perceptions compel the will to assent.¹⁸ For example, describing his *cogito* insight, Descartes remarks: “. . . I *could not but judge* that something which I understood so clearly was true” (AT 7:58 / CSM 2:41; our emphasis). In the Fifth Meditation he says more generally: “my nature is such that so long as I perceive something very clearly and distinctly I cannot but believe it to be true” (AT 7:69 / CSM 2:48).¹⁹

However, when “the intellect does not have sufficiently clear knowledge,” the will

is indifferent as to whether [the agent] should assert or deny either alternative, or indeed refrain from making any judgment on the matter . . . For although probable conjectures may pull me in one direction, the mere knowledge that they are simply conjectures, and not certain and indubitable reasons, is itself quite enough to push my assent the other way. (AT 7:59 / CSM 2:41)

Indifference is a state of motivational ambivalence in the will. (AT 7:58 / CSM 2:40; AT 4:173 / CSMK 245). It gives us “freedom to assent or not to assent” to what the intellect (imperfectly) displays (AT 7:61 / CSM 2:42).

What should we do with this freedom? Descartes answers:

If . . . I simply refrain from making a judgement in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly and avoiding error. But if in such cases I either affirm or deny, then I am not using my free will correctly . . . since *it is clear by the natural light that the perception of the intellect should always precede the determination of the will.* (AT 7:59-60 / CSM 2:41; our emphasis)

¹⁸ As will be discussed below (end of section II.D), some interpreters would quarrel with this characterization if it is taken to imply a sort of intellectual determinism by C&D perceptions; however, it is not controversial that Descartes thought C&D perceptions were compelling *in some sense*, as opposed to non-compelling non-C&D perceptions.

¹⁹ Descartes repeats many times elsewhere in the *Meditations* and in other places the idea that he cannot but believe what he clearly and distinctly perceives (see AT 8a:21 / CSM 1:207; AT 4:116 / CSMK233; AT 7:145 / CSM 2:104; AT 7:166 / CSM 2:117; AT 7:432-3 / CSM 2:292).

The final italicized line expresses the rule for correct judgment:²⁰

Norm of Assent: Assent only to what is clearly and distinctly perceived.²¹

When we extend the use of the will “to matters which [we] do not understand” (AT 7:58 / CSM 2:40), we knowingly violate this norm and *fail* to use the will properly. Thus Descartes says:

In this incorrect use of free will may be found the privation which constitutes the essence of error. The privation, I say, lies in the operation of the will in so far as it proceeds from me, but not in the faculty of will which I received from God, nor even in its operation, in so far as it depends on him. (AT 7:60 / CSM 2:41)

This is the core of Descartes’ defense. God is a perfect manufacturer who has designed and built our belief-production faculties to generate true beliefs. This implies a divine warrantee on the truth of our beliefs, but only on those that are formed *in accordance with manufacturer instructions*—i.e. those formed according to the Norm of Assent. Error is still possible because we are free to flout those instructions. Hence, Infallibility does not follow from the Divine Guarantee (or its theological grounds).

The Norm of Assent also entails Correctness. If I should assent only to what is C&D, then when I do so assent, I am using my faculty of judgment correctly. Therefore, the Norm of Assent appears to be central not only to Descartes’ defensive project, but also to his derivation of the Truth Rule. Without Correctness, he cannot reach the Truth Rule, and Correctness seems to rest, in turn, on the Norm of Assent.

II.B. The Appearance of Circularity

²⁰ In the italicized line, “perception of the intellect” must refer to *clear and distinct* perception. Otherwise, it would be permissible to pass judgment on anything, for *some sort* of perception is a prerequisite for judgment (AT 8b:363 / CSM 1:307; AT 7:377 / CSM 2:259).

²¹ In *Principles* I.44, Descartes puts the rule thus: “When we give our assent to something which is not clearly perceived, this is always a misuse of our judgement . . . since the light of nature tells us that we should never make a judgment except about things we know” (AT 8A:21 / CSM 1:207). Della Rocca (2006, 157) offers the following gloss: “[W]e should assent to clear and distinct ideas and should not assent to non-clear and distinct.”

Michael Della Rocca points out an apparent circularity in Descartes' argument by asking: "Why, for Descartes, should we assent only to clear and distinct ideas?"²² Della Rocca finds Descartes' answer in the following parallel texts from the *Meditations* and *Principles*:²³

If . . . I simply refrain from making a judgement in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly and avoiding error. But if in such cases I either affirm or deny, then I am not using my free will correctly. *If I go for the alternative which is false, then obviously I shall be in error; if I take the other side, then it is by pure chance that I arrive at the truth.* (AT 7:59-60 / CSM 2:41; our emphasis)

"When we give our assent to something which is not clearly perceived, *this is always a misuse of our judgment . . . either we go wrong, or if we do stumble on the truth, it is merely by accident, so that we cannot be sure that we are not in error.* (AT 8A:21 / CSM 1:207)

Focusing on the italicized sections of these passages, Della Rocca concludes:

Descartes seems to be saying that it is because in assenting to non-clear and distinct ideas our judgment is at best accidentally true and at worst false that we should not assent to such ideas. Rather, we should restrict our assent to ideas—clear and distinct ideas—that are not true only by chance, that are guaranteed to be true.²⁴

²² Della Rocca, "Judgment and Will," 157. We are not completely convinced that Descartes has an answer to this question. He may think that NA is simply self-evident. However, throughout this paper we will assume with Della Rocca that Descartes at least implicitly has an argument for NA.

²³ Della Rocca also appeals to texts from the Second Replies (AT 7:144 / CSM 2:102-3) and Sixth Meditation (AT 7:80 / CSM 2:55-56) where Descartes says that clear and distinct ideas can correct our other ideas. He further points to the italicized line in the following passage from the Fourth Meditation: "since my understanding comes from God, *everything that I understand I undoubtedly understand correctly, and any error here is impossible*" (our emphasis; AT 7:58 / CSM 2:40). We grant that in context, these passages probably assert that clear and distinct perceptions are true. However, none of these passages explains Descartes' reasons for holding the norm of assent. Therefore, they fall short of what Della Rocca needs to substantiate his interpretation.

²⁴ "Judgment and Will," 157.

But if this is correct, then Descartes derives the Norm of assent from the thought that “. . . clear and distinct ideas are guaranteed to be true and non-clear and distinct are not.”²⁵ But this latter thought is a *strengthened* version of the truth rule, a “Super Truth Rule.”²⁶

So the Truth Rule depends on Correctness, which depends on the Norm of Assent, which in turn depends on the (Super) Truth Rule. This obviously begs the question. As Della Rocca explains,

One of the premises needed for arguing that clear and distinct ideas are true is . . . that we should assent only to clear and distinct ideas. This claim in turn requires argument, and the argument for it seems to be based on the claim that clear and distinct ideas are guaranteed to be true. So, one of the premises for the claim that clear and distinct ideas are true is that clear and distinct ideas are guaranteed to be true. Here the conclusion—indeed, a strengthened version of the conclusion—is itself a premise in the argument.²⁷

This “internal” circle would be a problem for Descartes even in readings that avoid the traditional “external” circle. Suppose that Descartes is, at bottom, just trying to show that reason can remove skeptical doubts. If removing doubt is an exercise of *reason* and not sheer will, then it must follow rational procedures. And begging the question is not a rational procedure.

II.C. An Unsuccessful Way Around the Problem

But perhaps Della Rocca has misunderstood Descartes. Perhaps his argument for the Truth Rule does not really depend on Correctness. In the Fourth Meditation, Descartes does not explicitly state the Truth Rule until the final paragraph, at which point he provides the following argument for it:

. . . [I]f, whenever I have to make a judgement, I restrain my will so that it extends to what the intellect clearly and distinctly reveals, and no further, then it is quite impossible for me to go wrong. *This is because* [II] every clear and distinct perception

²⁵ Ibid.

²⁶ It is strengthened because it claims not just that all clear and distinct ideas are true, but that they are *guaranteed* to be true, and *only* they are so guaranteed.

²⁷ Della Rocca, “Judgment and Will,” 158-59.

is undoubtedly something, and hence cannot come from nothing, but must necessarily have God for its author. Its author, I say, is God, who is supremely perfect, and who cannot be a deceiver on pain of contradiction; hence the perception is undoubtedly true. (AT 7:62 / CSM 2:43)

Many commentators consider *this* to be Descartes' argument for the truth rule, and as John Carriero has noted, it "does not obviously depend on any claims he has made about how things should or must go with my will" or "on his views about how I should assent."²⁸

To put the point even more strongly, this passage seems to contain the following argument:

- A) Every clear and distinct perception is something.
- B) [Whatever is something comes from God.]
- C) Every clear and distinct perception comes from God. [from A & B]
- D) God is not a deceiver.
- ∴\ Every clear and distinct perception is true. [from C & D]

Premises C and D resemble the principles of divine origin and divine veracity, but Descartes uses them here to argue *directly* for the truth rule. This reasoning manifestly does *not* depend on Correctness or the Norm of Assent.

It may be, then, that Correctness and the Norm of Assent are merely parts of Descartes' free will defense. To employ them for that hypothetical defensive task, he need not commit to their truth, nor to the truth of their grounds. Therefore, even if the Super Truth Rule is among those grounds, Descartes does not need it to be actually true. Once he has used it to restore trust in his theological premises, Descartes can throw it away (along with the other elements of his free-will defense) before moving on to prove the Truth Rule in the Fourth Meditation's final paragraph. Hence, in this reading there is no vicious circle to be found.

Although the reading just sketched would exonerate Descartes of Della Rocca's charges against him, we do not find this plausible. For this reading sharply separates the argument of the final paragraph from the free-will defense, when in fact they are connected. The argument of the final

²⁸ Carriero 2009: 463, note 40.

paragraph depends on ideas introduced in the earlier account of the will. This is evident both from considerations internal to the final paragraph and from other significant textual evidence.

If the final paragraph does not draw on the free will defense, the argument there is invalid. In stating his conclusion, Descartes explicitly mentions both judgment and will: “if, whenever I have to make a judgment, I restrain my will so that it extends to what the intellect clearly and distinctly reveals, and no further, then it is quite impossible for me to go wrong.”²⁹ But read in isolation from the free will defense, the premises A-D warrant a conclusion only about C&D perceptions *in the intellect*. They entail nothing about judgments or beliefs, making the argument invalid.

Furthermore, Descartes invokes the free-will defense in the final paragraph when he says: “every clear and distinct perception is undoubtedly (*proculdubio*) something.” “Undoubtedly” insinuates his claim that C&D perceptions leave no indifference in the will, no room for doubt. We cannot but assent to such perceptions, so they *always* issue in judgment. Given this implicit assumption, when Descartes concludes that “the perception is undoubtedly true,” (AT 7:62 / CSM 2:43) he means that the perception *and its accompanying judgment* together constitute a true belief.³⁰ When read as relying on the free will defense in this way, the argument of the final paragraph is valid.

The validity issues just discussed suggest that the Fourth Meditation’s final paragraph is probably a shorthand version of this longer argument for the Truth Rule from the Second Replies:

[*Divine Veracity*:] God . . . must . . . be supremely good and true, and it would therefore be a contradiction that anything should be created by him which positively tends towards falsehood. [*Divine Origin*:] Now everything real which is in us must have been bestowed on us by God . . . moreover, we have a real faculty for recognizing the truth and distinguishing it from falsehood . . . [*Divine Guarantee*:] Hence this faculty must tend towards the truth, at least when we use it correctly [*Correctness*:] (that is, by assenting only to what we clearly and distinctly perceive, for no other correct method of employing this faculty can be imagined). For if it did not so tend, then, since God

²⁹ He *should* mention them, for the following reasons. Descartes wants to be sure that when he believes what he C&D perceives, his belief is true. Descartes defines truth as “conformity of thought with its object” (AT 2:597 / CSMK 139) and in the Third Meditation he suggests that mistakes or errors involve wrongly “judging that the ideas which are in me resemble, or conform to, things located outside me” (AT 7:37 / CSM 2:26). By parity, a true belief would be a *correct* judgment that one of my ideas conforms to some external object. And as a *judgment*, a true belief requires the activity of both intellect and will.

³⁰ In other words, “the perception” here means more than just the bare content displayed in the intellect alone.

gave it to us, he would rightly have to be regarded as a deceiver . . . And since it is impossible to imagine that he is a deceiver, [*Truth Rule*.] whatever we clearly and distinctly perceive must be completely accepted as true and certain. (AT 7:144 / CSM 2:103)

Here, as our editorial insertions highlight, Descartes argues for the Truth Rule by means of two syllogisms—just as in our reconstruction above (pp 6-8). First, he argues from Divine Veracity and our Divine Origin to the Divine Guarantee: *necessarily, we arrive at truth when judging correctly*. Then he combines this Divine Guarantee with Correctness—*we judge correctly when we assent to C&D perceptions*—to infer the Truth Rule: *when we assent to C&D perceptions, we arrive at Truth*.

Because Descartes' argument for the Truth Rule uses Correctness as a premise, Descartes must affirm the truth of Correctness, as well as its grounds. He cannot employ it only for hypothetical, defensive purposes. Thus far, Della Rocca's reading seems correct.

II.D. First Solution: Correctness and Compelled Assent

For Della Rocca's circularity charge to stick, Descartes must *derive* Correctness from the Norm of Assent. But though the Norm of Assent entails Correctness, it might not be Descartes' *grounds* for Correctness. For there is another part of Descartes' free will defense that also entails Correctness, and so could serve as its grounds. This other part is Descartes' doctrine that C&D perception compels assent.

Descartes clearly indicates that it is our *nature* to assent (without fail) to what we C&D perceive. We were made this way. Therefore, when we assent to clear perceptions, we act in accord with God's designed plan for us. But to use a faculty correctly *just is* to use it in the way that God intended for us to use it. Therefore, when we assent to C&D perceptions, we judge correctly. Indeed, because we cannot but assent to C&D perceptions, it is impossible for us *not* to act correctly in such cases.³¹

In this interpretation, the Truth Rule depends on Correctness, but Correctness depends on the compelling character of C&D perceptions and *not* on the Norm of Assent. The Norm of Assent

³¹ Carriero may be gesturing at something like this argument in the following passage: “. . . in order to establish the truth rule via a claim about assent and clear perception, it seems to me that all that Descartes really needs is the claim that I *must* assent to something that I perceive very clearly (regardless of how we sort out the question about whether I *should* assent).” (Carriero 2009: 463, note 40)

is not a premise in the argument for the Truth Rule. Therefore, even if the Norm of Assent depends on the Super Truth Rule, there is no vicious circle.

We think it plausible that Descartes grounded Correctness, and hence the Truth Rule, on his doctrine of compelled assent. However, the doctrine of compelled assent is controversial. Some interpreters think that nothing, not even C&D perceptions, can truly compel the will.³² Therefore, in the next and final section, we will show that circularity is an illusion even if Descartes grounds Correctness on NA.

II.E. A Second Solution: Two Kind of Chanciness

Suppose, then, that the truth rule depends on the Norm of Assent. According to Della Rocca, Descartes derives that norm from the following two more fundamental assumptions:

“Take No Chances” (TNC): “we should restrict our assent to ideas . . . that are not true only by chance, that are guaranteed to be true” (2011, 96)

and

Super Truth Rule (STR): ‘[O]nly clear and distinct ideas...are guaranteed to be true’ (2006, 157).

Given this derivation, the Truth Rule turns out to depend on the actual truth of STR, a strengthened version of itself.

However, the dependence of the truth rule (“clear and distinct ideas are true” (Della Rocca 2006, 158)) on STR is viciously circular only if the sense of “guaranteed truth” in STR is the *same* as the sense of “true” in the Truth Rule. But it quite clearly *cannot be* the same, as is evident from drawing a simple distinction and paying attention to the force of TNC as Descartes articulates it.

The distinction is between two different senses of “chanciness” relevant to Descartes’ project. The truth rule asserts a necessary connection between mind and world, such that we clearly and distinctly perceive *P* only if *P*; this is “objective non-chanciness.”³³ Objective chanciness, then, would

³² See, for example, Alquié 1966: 280-299, Alanen 1999, and Moyal 1996.

³³ Readers concerned that this sort of necessary connection pushes Descartes into strong validation should revisit the final paragraphs of section I. Descartes can distinguish between objective and subjective chanciness simply by appealing to the *ideas* of mind and world, even if his argument only reaches conclusions about what should be happening inside the mind.

be the absence of this connection, the possibility that we might C&D-perceive *P* even though it is not the case. This is the possibility invoked in the deceiving God scenario. Subjective chanciness, by contrast, is a matter not of a mind-world connection, but of how things stand *inside* the mind. If we are aware of a reason to doubt *P*, then *P* is subjectively chancy for us: with respect to *P*, “although probable conjectures may pull me in one direction, the mere knowledge that they are simply conjectures, and *not certain and indubitable reasons*, is itself quite enough to push my assent the other way” (AT 7:59 / CSM 2:41). Something is subjectively non-chancy (subjectively guaranteed) for us when we cannot think of a way for it to turn out false.

Descartes’ “take no chances” rule is about subjective chanciness. Arguably, Descartes articulates it in the second paragraph of the First Meditation when he says:

Reason now leads me to think that I should hold back my assent from opinions which are not completely certain and *indubitable* just as carefully as I do from those which are patently false. So, for the purpose of rejecting all my opinions, it will be enough if I find in each of them at least *some reason for doubt*. (AT 7:18 / CSM 2:12; our emphasis)

This rule instructs us to assent only when we have no reasons for doubt.

That TNC concerns subjective and not objective chanciness is further confirmed by Descartes’ attempts to follow it as the *Meditations* unfold. Prior to proving the truth rule, in an effort to “take no chances,” Descartes withholds assent from claims that he had previously clearly perceived, because he can now think of reasons to doubt them.³⁴ In fact—as he later establishes by proving the truth rule—these previously clearly-perceived claims are objectively non-chancy (guaranteed to be true). If TNC enjoined us to assent only to what is *objectively* guaranteed, then Descartes could have assented to the previously clearly-perceived claims. But he felt an obligation to withhold assent from them. This proves that TNC enjoins avoiding subjective rather than objective chanciness; it instructs us to give assent only to ideas that are “guaranteed” in the subjective sense of being indubitable.

But If TNC says “assent only to what is subjectively guaranteed,” and STR says “only clear and distinct ideas are objectively guaranteed,” they fail to share a common predicate and so cannot

³⁴ These previously clearly-perceived ideas are a case in which one perceives *P* only if *P* (there is an objective guarantee), but one can still think of reasons to doubt *P* (so that it is subjectively chancy). In such a case, Descartes thinks that one should withhold assent from *P*.

entail NA. Hence, on pain of invalidity, Descartes must intend STR to concern subjective chanciness, reasoning as follows:

- 1) TNC: We should assent only to what is subjectively guaranteed (we cannot doubt it).
 - 2) STR: Only clear and distinct ideas are subjectively guaranteed (we cannot doubt them).
- Therefore,
- 3) Norm of Assent: we should assent only to clear and distinct ideas.

Della Rocca's assertion that Descartes grounds the Norm of Assent on STR can remain plausible only when STR is construed as a *subjective* guarantee (an assertion that we can see *no reasons to doubt* what is clear and distinct).

Descartes' procedure in this reading must be to first establish the Norm of Assent by reasoning about *subjective* chanciness (from TNC and STR), and then to use that norm to argue further that clear and distinct perception is not *objectively* chancy. Since God did not give him a faculty that would allow him to err when used correctly, and assenting to clear perceptions (while thus perceiving them) is correct, there is a necessary connection between his clear and distinct perceptions and the world they represent—i.e., the truth rule is established. This argument is not viciously circular because the conclusion concerns a *different sense* of non-chanciness (or guarantee) than does his initial assumption: Descartes assumes a *subjective* guarantee and proves (in the end), an *objective* guarantee. Even granting Della Rocca's interpretation, then, vicious circularity does not follow.

Conclusion

We have shown, first of all, that Descartes' Fourth Meditation argument for the Truth Rule does not, by its very existence, rule out non-circular readings of the Third Meditation. The Fourth Meditation can be understood as simply completing the removal of doubt begun in the Third Meditation. Therefore, it provides no evidence for the traditional "external" Cartesian Circle (we take no stand on whether there is enough evidence from other texts to motivate this problem). We have also shown that despite initial appearances, there is no "internal circle" or *petitio principia* in the Fourth Meditation itself. In accomplishing this second task, we have looked carefully at Descartes' argument for the Truth Rule, showing it to stretch throughout the Fourth Meditation and to depend crucially on the claim that it is correct to assent to C&D ideas (Correctness). Correctness itself may depend

either on Descartes' doctrine of compelled assent, or—as Della Rocca contends—on the Norm of Assent. But either way, Descartes' argument is not viciously circular.

Works Cited

- Alanen, Lilli K. (1999) 'Intuition, Assent and Necessity: The Question of Descartes's Psychologism', *Acta Philosophica Fennica*, 64, 99-121.
- Alquié, Ferdinand (1966) *La découverte métaphysique de l'homme chez Descartes* 2nd ed. Paris: Presses Universitaires de France.
- Augustine and Williams, Thomas (1993) *On Free Choice of the Will*. Hackett Pub Co Inc.
- Carriero, John Peter (2009) *Between two worlds : a reading of Descartes's Meditations*. Princeton: Princeton University Press.
- Cress, Donald (1994) 'Truth, Error, and the order of reasons: Descartes' puzzling synopsis of the fourth meditation' in Cottingham, ed. *Reason, Will, and Sensation: Studies in Descartes's Metaphysics*. Oxford: Clarendon Press.
- Della Rocca, Michael (2006) 'Judgment and Will' in Gaukroger, ed. *The Blackwell Companion to Descartes' Meditations*.
- Della Rocca, Michael (2011) 'Taking the Fourth: Steps toward a New (Old) Reading of Descartes', *Midwest Studies In Philosophy*, 35(1), 93-110.
- Dicker, Georges (1993) *Descartes: An Analytical and Historical Introduction*. Oxford University Press, USA.
- Fulmer, Everett and C.P. Ragland (2017) 'Against the New Cartesian Circle,' *Canadian Journal of Philosophy* 47 (1): 66-74.
- Hatfield, Gary (2014) *The Routledge Guidebook to Descartes' Meditations*. London and New York: Routledge.
- Loeb, Louis E. (1990) 'The Priority of Reason in Descartes', *Philosophical Review*, 99(1), 3-43.
- Moyal, Georges J. D. (1996) 'Magicians, Doubters and Perverts', *Revue Internationale de Philosophie*, 50(195), 73-107.
- Newman, Lex (1999) 'The Fourth Meditation', *Philosophy and Phenomenological Research*, 59(3), 559-591.
- Plantinga, Alvin (1990) 'God, Evil, and the Metaphysics of Freedom' in Adams and Adams, eds., *The Problem of Evil*. Oxford: Oxford University Press, 83-109.
- Ragland, C. P. (2006) 'Is Descartes a Libertarian?' in Garber and Nadler, eds., *Oxford Studies in Early Modern Philosophy*. Oxford: Oxford University Press, 57-90.

Ragland, C. P. (2016) *The Will to Reason: Theodicy and Freedom in Descartes*. Oxford University Press, USA.

Wagner, Stephen I. (2014) *Squaring the circle in Descartes' Meditations : the strong validation of reason*.
Cambridge University Press.