

# Against the new Cartesian Circle

Everett Fulmer and C.P. Ragland

Philosophy Department, Saint Louis University, Saint Louis, MO, USA

## ABSTRACT

In two recent papers, Michael Della Rocca accuses Descartes of reasoning circularly in the Fourth Meditation. This alleged new circle is distinct from, and more vicious than, the traditional Cartesian Circle arising in the Third Meditation. We explain Della Rocca's reasons for this accusation, showing that his argument is invalid.

**ARTICLE HISTORY** Received 11 December 2015; Accepted 8 August 2016

**KEYWORDS** Descartes; meditations; circularity; Cartesian Circle; fourth meditation; Michael Della Rocca

## 1. Introduction

In two recent papers, Michael Della Rocca locates a distinct instance of circular reasoning in Descartes' Fourth Meditation – i.e. distinct from the traditional 'Cartesian Circle' of the Third Meditation (2011, 2006).<sup>1</sup> This new circle, Della Rocca alleges, is an even more intractable fallacy: it is a bona fide instance of premise circularity in which the 'truth rule' ('everything I clearly and distinctly perceive is true') functions as a sub-premise in its own proof. While Della Rocca's provocative reading has garnered substantial attention in the literature (see, e.g. Carriero 2009, 278–9, 462–3; Christofidou 2013; 245 n. 10; Naaman-Zauderer 2010, 98–100), it has not yet received the refutation it deserves. We supply that refutation here. We show that even granting Della Rocca's interpretation of the Fourth Meditation, his charge of circularity is unwarranted.

## 2. The problem of meditation IV

Della Rocca begins his interpretation of the Fourth Meditation by noticing that it is structurally analogous to the Third Meditation. He writes,

---

**CONTACT** Everett Fulmer  [efulmer@slu.edu](mailto:efulmer@slu.edu)

© 2016 Informa UK Limited, trading as Taylor & Francis Group

After [Descartes] lays 'it down as a general rule' near the beginning of Meditation III 'that whatever I perceive very clearly and distinctly is true' [AT VII 35, CSM II 24]<sup>2</sup>, Descartes immediately casts this rule into doubt by raising the possibility that God is a deceiver. So, too, after finally offering an argument later in Meditation III for the claim that God is not a deceiver, Descartes immediately goes on in Meditation IV to raise a doubt about this claim. (2006, 142–3, 2011, 94)

Descartes raises this worry about God's non-deceptiveness in the following passage:

[S]ince God does not wish to deceive me, he surely did not give me the kind of faculty which would ever enable me to go wrong while using it correctly. There would be no further doubt [*dubium*] on this issue 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 VII 54, CSM II 37–8)

Della Rocca glosses the problem here as follows:

But obviously, as Descartes recognizes, I do sometimes make mistakes, and so the argument for the claim that God is not a deceiver and that clear and distinct ideas are true is called into doubt. Descartes will not have removed this doubt and beaten back the skeptic about clear and distinct ideas until he has shown how it is compatible with God's nondeceptiveness that we sometimes err. (2006, 143, 2011, 94)

To resolve this problem, Descartes must 'answer the following question: How, if God is not a deceiver and thus does not allow clear and distinct ideas to be false, can God allow that I ever make mistakes, in particular with regard to ideas that are not clear and distinct?' (2006, 143, 2011, 94). In other words, Descartes needs to show that while God's veracity ('God is not a deceiver') entails the truth rule ('clear and distinct ideas are true'), it does not entail human infallibility.<sup>3</sup>

Descartes blocks the inference to human infallibility with a 'freewill defense' that appeals to a distinction between 'user-error' and 'manufacturer-error'<sup>4</sup> along with a norm of correct assent. Descartes reasons that if we judge correctly and go wrong, then our faculty of judgment must have a design or manufacturing problem, for which God would be culpable. But if we ignore our manufacturer's instructions for correct use – if the cause of error lies not in the faculties themselves but in our misuse of those faculties – then God is not to blame. Descartes thinks God has given us the following clear instruction for correctly using our faculty of judgment:

[I]t is clear by the natural light that the perception of the intellect should always precede the determination of the will. (AT VII 60, CSM II 41)

Della Rocca paraphrases this deliverance of the natural light as follows:

*Norm of Assent (NA):* "[W]e should assent only to clear and distinct ideas". (2006, 157)<sup>5</sup>

We are free to follow NA or to flout it. When we flout it, we place ourselves outside of the divine guarantee by using our faculties contrary to the manner God

intends. The result is a ‘user-error’ for which the ‘manufacturer’ of our faculties bears no responsibility. Thus, God’s goodness does not entail total infallibility, but only infallibility *when we judge correctly*.

### 3. The case for the new circle

Della Rocca’s circularity charge begins by asking Descartes to justify the norm of assent. Della Rocca asks: ‘Why, for Descartes, should we assent only to clear and distinct ideas? Without a good reason for this claim, Descartes will lack an effective way of putting to rest his doubt about clear and distinct ideas’ (2006, 157). Della Rocca then contends that Descartes realized this need and attempted to justify NA in the following passage:

But if in such [non-clear and distinct] 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 VII 59–60, CSM II 41)

According to Della Rocca, Descartes here affirms that we should not assent to *non-clear-and-distinct* perceptions because such judgments would be either false or ‘at best accidentally true’ (2006, 157). Conversely, clear and distinct ideas are neither false nor accidentally true. From which it follows<sup>6</sup> that clear and distinct ideas are true and guaranteed to be so. Thus, Della Rocca concludes:

The reason Descartes offers [to justify the norm of assent] seems to be ... [that] clear and distinct ideas are guaranteed to be true and non-clear and distinct are not. (2006, 157)

But offering this justification for NA is a fatal error. The truth rule, recall, says that all clear and distinct ideas are true. But here Descartes is appealing to an even stronger notion – not only are all clear and distinct ideas true, but they are guaranteed to be so and only they are so guaranteed. We will refer to this strengthened version of the truth rule as the

*Super Truth Rule* (STR): “[O]nly clear and distinct ideas ... are guaranteed to be true”. (Della Rocca 2006, 157)

According to Della Rocca, Descartes derives NA from STR. But given the purported aim of Meditation IV, grounding the norm of assent in this way 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. (2006, 158–9)

And therefore, Della Rocca concludes, Descartes is guilty of a particularly vicious form of circular reasoning in the Fourth Meditation: here, the truth rule is not just prematurely trusted – it is used as a premise to prove itself.

#### 4. Our initial refutation

We begin with a fairly simple argument. Notice the structure of the problem Descartes faces: the goodness of God and the proof of the truth rule are called into question *because* God's veracity (G) *seems* to entail the outlandish claim that humans are totally infallible (I). To solve this problem, Descartes merely needs to show that (I) does *not* follow from (G). And since (I) is inconsistent with the actuality of human error (E), showing that (G) does not entail (I) is equivalent to showing that the following set is consistent: {(G), (E)}.

Furthermore, demonstrating the consistency of a set of propositions *only* requires appeal to the logical relations that hold among the members of that set. It does not require any additional factual claims. That is to say, the entire discourse of such a project is non-committal about the truth of the notions it employs. Descartes need only invoke elements of his 'freewill defense' – including NA and whatever grounds it – as hypotheses for the purposes of showing that (E) is consistent with (G). Even if Descartes does appeal to STR to ground NA, such an appeal does not commit him to the truth of STR. But, if Descartes does not need to endorse the truth of STR, then his use of STR does not presuppose the final conclusion of his argument – the truth rule. Thus, there is no circle.

#### 5. An analogical rejoinder

In Della Rocca's second paper (2011), he responds to something like this objection: '[o]ne might argue,' he notes, that 'all Descartes needs to do is to show that God's veracity, and thus the guaranteed truth of clear and distinct ideas, is consistent with our occasional errors with regard to ideas that are not clear and distinct' (99). However, he quickly retorts: 'But again, the point is: How does Descartes show this consistency?' (99). He does so, Della Rocca insists, by illegitimately and prematurely relying on the claim that clear and distinct ideas are guaranteed to be true. Della Rocca makes his point by way of the following analogy:

Imagine trying to reconcile one's theory that Joe is the murderer with apparently conflicting DNA evidence that points to someone else, Susie. You reconcile these by claiming that the DNA evidence is misleading. But, let us say, your only basis for claiming that the DNA evidence is misleading is that Joe (and not Susie) is the murderer ... In this case, you have not really defended your theory in light of the DNA evidence. What you need is a way of discounting the apparent counterevidence, a way that is independent of the claim that Joe is the murderer ... To avoid circularity, Descartes needs an independent way of reconciling occasional errors, on the one hand, with the guaranteed truth of clear and distinct ideas and God's veracity, on the other. (2011, 99)

Despite its intuitive pull, Della Rocca's Joe and Susie analogy misrepresents Descartes' dialectical situation. The analogy is based on the inductive, probabilistic reasoning of courtrooms. If the attorney appeals to 'Joe is the murderer' in response to his opponent, he must assume the actual truth of that claim and so beg the question. But the same does not hold for Descartes since the problem Descartes faces is *not* inductive. What Descartes must overcome is that the reasoning that leads to the truth rule seems also to *entail* human infallibility (I). That is to say, there is a purported logically necessary connection between God's veracity (G) and (I).<sup>7</sup> Hence, Descartes' dialectical situation differs considerably from the attorney seeking to prosecute Joe.

Once one sees the switch of dialectical contexts, the analogy loses its force. Given that Descartes must only demonstrate the logical consistency of the set  $\{(G), (E)\}$ , he must do no more than show (semantically speaking) that there is at least one logically possible truth value assignment that would make all statements of the set true. Descartes need not say anything about the actual truth values of the statements he invokes. And since the super truth rule is a logical possibility, there is nothing illegitimate about assuming it. Therefore, even if Descartes grounds NA on STR for the purposes of blocking the inference to (I), he does not beg the question thereby.

## 6. A modal rejoinder

Della Rocca has responded to our last line of reasoning by noting that on the most commonly used modal logic – System S5, which he believes Descartes would accept – the possibility of the super truth rule entails the actuality of the truth rule (2014).<sup>8</sup> And thus even given our defense above, Descartes would still be guilty of circularity.

To see Della Rocca's rejoinder, consider that the truth rule (TR) says that all clear and distinct ideas are true, and the super truth rule (STR) says that all (and only) clear and distinct ideas are *guaranteed* to be true (N.B. guaranteed = necessary, for Della Rocca). Therefore, STR asserts the necessity of TR ( $\Box$ TR). So, if Descartes assumes the mere possibility of the super truth rule, he thereby assumes:

$$\Diamond\Box\text{TR}$$

This means that at some accessible possible world – call it world 2 – the truth rule is necessarily true, i.e.  $(\Box\text{TR}, w_2) = \text{T}$ . Now, System S5 has Euclidean frame conditions, which entails that accessibility relations are symmetric. Thus, since world 1 (the actual world) accesses world 2, it follows that world 2 must also access world 1. And thus, since  $\Box\text{TR}$  is true at world 2 it would follow by the valuation rule for ' $\Box$ ' that TR is true at the actual world. So on System S5,

$$\Diamond\Box\text{TR} \vdash \text{TR}$$

Hence, Della Rocca concludes, even if Descartes merely assumes the super truth rule as a possibility, he thereby assumes the actual truth of the truth rule, and thus begs the question.

We wonder about attributing symmetric frame conditions to Descartes. With his creation doctrine of eternal truths, Descartes seems to posit two levels of modal truths – absolute and restricted to creation – that do not access each other (see, e.g. [AT I 152, CSMK 25]; [ATV 224, CSMK 358–9]; and [AT IV 118, CSMK 235]). And this would imply non-symmetric frame conditions.<sup>9</sup> But even if we set that thorny issue aside, Della Rocca's objection is still flawed in a more straightforward sense. With or without S5, the way in which Descartes must regard STR, for purposes of showing set consistency, is not accurately represented as a full-blooded modal commitment to the possibility of STR in the actual world.

Notice that whether or not a set is consistent is wholly a matter of the terms that are *already* members of the set. Sets can be made inconsistent by the addition of new members, but inconsistent sets *cannot* be made consistent via addition. The set  $\{p, \sim p, q\}$  will always remain inconsistent no matter how many terms are added. Thus, when Descartes demonstrates to the reader that error is consistent with his premises, he does not, and cannot, *make* the set consistent, if it is not already so. For Descartes to succeed, the set must already be consistent.

But if the set is already consistent, then what is Descartes doing with his story about our ability to flout the norm of assent? At least in terms of the project to 'block' the apparent entailment to (I), all such talk can only be pedagogical. Descartes is merely painting a conceptual backdrop that makes salient to the reader those aspects of the set  $\{(G), (E)\}$  that render it consistent.<sup>10</sup> All features of such a hypothetical story can be fully disposed of once they achieve their purpose – i.e. once the reader sees that the set is consistent.

Moreover, the story Della Rocca attributes to Descartes (including STR) is not the only one available for this purpose. A story without STR could work just as well. For example, Descartes could say that humans are unable to withhold assent to clear and distinct perceptions, but *are* able to withhold it from non-clear and distinct perceptions.<sup>11</sup> Given that 'ought implies can,' this difference in our abilities would explain why NA obliges us to withhold assent from non-clear and distinct ideas but not from clear and distinct ones. This alternative grounding for NA, along with other features of Descartes' freewill defense, would show that (G) and (E) can both be simultaneously true.<sup>12</sup>

Hence, if Descartes chooses to use STR as part of his hypothetical story, this is merely a rhetorical choice.<sup>13</sup> By using such a pedagogical aid, Descartes does not commit himself to any metaphysical truths at all, not even to the truth of any possibility claims. Therefore, even if symmetric frame conditions were appropriate for Cartesian modalities (see n. 9), there would still be nothing viciously circular about including STR in his hypothetical story.

## 7. Conclusion

Michael Della Rocca's 'new circle' is based on several interpretive claims: (1) while attempting to prove the truth rule, Descartes realizes that his premises seem to lead to a patently false result – namely that we are infallible. (2) Descartes must first solve this puzzle in order to continue his proof of the truth rule. (3) In the course of solving this puzzle, Descartes relies on a norm of correct assent (NA). And, (4) Descartes grounds NA on the super truth rule (STR) – all and only clear and distinct ideas are guaranteed to be true. But, (5) STR is just a strengthened version of the truth rule. Thus, Della Rocca infers that by grounding NA in this way, Descartes is guilty of premise circularity – he literally uses the truth rule to prove itself.

In the refutation above we have intentionally avoided contesting any of Della Rocca's core interpretive claims (1) through (5). Instead, we have shown that even granting Della Rocca his reading of Meditation IV, it does not follow that Descartes is guilty of premise circularity. The reason, in short, is that the project to show that an entailment does not hold is merely the project to show a given set is consistent. And this project *only* requires true claims about the logical relations that hold among the members of the set. It requires *no* commitment to truths about reality beyond. Hence, both Della Rocca's Joe and Susie analogy and his modal rejoinder are misplaced, and, Della Rocca has failed to find a new instance of circularity in Meditation IV.

## Notes

1. The traditional Cartesian Circle concerns the fact that (i) Descartes suggests that he cannot fully trust the 'truth-rule' until he first knows that God exists. But (ii) it seems he cannot trust the premises of his proof for God's existence without first relying on the truth rule. Unlike Della Rocca's new circle, this is *not* a case of premise circularity.
2. References to Descartes refer to volume and page numbers from these standard editions:
  - AT Adam and Tannery, 1974–86.
  - CSM Cottingham, Stoothoff, and Murdoch 1984–5.
  - CSMK Cottingham et al. 1991.
3. One might wonder whether God's veracity alone entails the truth rule. We take up this question in 'The Fourth Meditation and Cartesian Circles' (unpublished manuscript). But for our purposes here, Della Rocca's reading is what matters. Note that in the passage just quoted, Della Rocca says 'God is not a deceiver and *thus* does not allow clear and distinct ideas to be false' (our emphasis). His belief in this entailment explains why Della Rocca takes divine veracity and the truth rule to both follow from one and the same (Third Meditation) argument, and it explains why he suggests that once Descartes has removed doubts about this argument, he will have 'beaten back the skeptic about clear and distinct ideas.'
4. We take this terminology from Newman (1999, 559–91).
5. On the same page, he states the same rule a bit more precisely: '[W]e should assent to clear and distinct ideas and should not assent to non-clear and distinct.'

6. N.B. it only follows if 'chance' is read as claim of objective, instead of subjective, probability. But we shall set this interpretive issue to one side (see: Carriero 2009, 462–3).
7. If one wonders about this interpretive claim, note that a probabilistic challenge could not throw Descartes' premises into doubt, for recall that in the Third Meditation he discovered a deductive proof for (G) (cf. Descartes' formal presentation of his argument AT VII 160–70, CSM II 113–20). From an inductive challenge, therefore, it would not follow that (G) is false, but only that something unlikely (error, in this case) had occurred. Thus, if there is a problem here at all, it must be a deductive problem. See Newman (1999) for another argument to the effect that the Fourth Meditation problem must be deductive.
8. Personal correspondence (May 24, 2014).
9. Della Rocca's own work on Cartesian modality seems to bear this out (see, e.g. 2005, 1–33, 2011, 103–5). Della Rocca argues that the two levels of modal discourse in Descartes *do not* have open accessibility relations to each other. But to deny universal accessibility entails denying symmetry. From the absolute perspective that considers the omnipotence of God, everything is contingent (or everything minus facts about God's own nature). Whereas, from the more restricted perspective of created reality certain truths are genuinely necessary – such as  $2 + 2 = 4$ . From the absolute perspective, God could have made a different eternal truth than  $2 + 2 = 4$ . He could have made  $2 + 2 = 4$  false and some other claim true such as  $2 + 2 = 5$ . Thus, in the absolute sense, there is a possible world where  $2 + 2 = 5$ . And in this possible world  $2 + 2 = 5$  would have been an eternal truth, that is to say, it would have been necessary in the restricted sense. Thus, if we read ' $\diamond$ ' as referring to absolute modality and ' $\square$ ' as referring to restricted modality, Descartes is saying that  $\diamond \square (2 + 2 = 5)$ . And this claim is true in the actual world. It is true here and now that God could have made different eternal truths. Yet, the possible world referred to in the expression  $\diamond \square (2 + 2 = 5)$  does not access the actual world as symmetry requires. For otherwise, we could infer that in the actual world both  $2 + 2 = 5$  and  $\square (2 + 2 = 4)$  are true, which would be a contradiction. Thus, we wager, symmetry cannot be attributed to Descartes.
10. Salient here, e.g. would be (i) the fact that divine benevolence does not entail deterministic divine control over every creaturely act, (ii) that manufacturer responsibility does not entail responsibility for user errors, and (iii) the goodness of the whole of reality does not entail that at a given time slice, for a particular sub-section of reality, improvement is impossible.
11. N.B. Some commentators think that this scenario is Descartes' actual considered view (see, e.g. Carriero 2009, 462 n. 40; Kenny 1972).
12. In fact, once the pedagogical nature of the story is seen, it becomes clear that Descartes need not justify NA at all for the purpose of showing the consistency of  $\{(G), (E)\}$ . For that project, NA could be nothing more than an arbitrarily chosen fiction and still succeed – at least if blocking (I) is the only argumentative project for which Descartes employs NA. It might be objected that Descartes needs NA for another project – that of positively establishing the truth rule. One might argue that Descartes cannot establish the truth rule without claiming that humans act blamelessly (i.e. use their faculty of judgment properly) when they assent to clear and distinct perceptions; but Descartes cannot establish that, in turn, without putting forward NA as actually true. While this suggestion is not implausible, there are reasons to doubt it. As Carriero (2009, 463 n. 40) notes, when Descartes finally states the truth rule in the last paragraph of the Fourth Meditation, his argument there does not explicitly depend on any claim about how one



should (or should not) assent. In ‘The Fourth Meditation and Cartesian Circles’ (unpublished manuscript), we explore the question of whether Descartes appeals to the actual truth of NA in establishing the truth rule. However, this question is not relevant to our refutation of Della Rocca, who gives every appearance of holding that the truth rule will follow immediately – without further appeal to NA – once Descartes has successfully demonstrated the consistency of {(G), (E)}. See note 3 above.

13. That is to say, if Della Rocca is correct then Descartes may be making a rhetorical blunder, but even so he is not making a logical one.
- \* Thanks to Michael Della Rocca, Lex Newman, the anonymous referees, and the audiences at both the 2015 APA Central Division Meeting and the 2014 Midwest Conference in Early Modern Philosophy (at UW-Milwaukee) for feedback on earlier versions of this paper.

## Notes on contributors

**Everett Fulmer** is a dissertation research fellow at Saint Louis University.

**C.P. Ragland** is an associate professor of philosophy at Saint Louis University. He is the co-editor of *What is Philosophy?* (Yale U.P., 2001) and author of *The Will to Reason: Theodicy and Freedom in Descartes* (Oxford U.P., 2016).

## References

- Adam, Charles and Paul Tannery, eds. 1974–86. *Oeuvres de Descartes* [Works of Descartes]. 2nd ed. 11 vols. Paris: Vrin/CNRS.
- Carriero, John. 2009. *Between Two Worlds: A Reading of Descartes's Meditations*. Princeton, NJ: Princeton University Press.
- Christofidou, Andrea. 2013. *Self, Reason, and Freedom: A New Light on Descartes' Metaphysics*. New York: Routledge.
- Cottingham, John, Robert Stoothoff, and Dugald Murdoch, trans. 1984–5. *The Philosophical Writings of Descartes*. 2 vols. Cambridge: Cambridge University Press.
- Cottingham, John, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny, trans. 1991. *The Philosophical Writings of Descartes: The Correspondence*. Vol. 3. Cambridge: Cambridge University Press.
- Della Rocca, Michael. 2005. “Descartes, the Cartesian Circle, and Epistemology Without God.” *Philosophy and Phenomenological Research* 70 (1): 1–33.
- Della Rocca, Michael. 2006. “Judgment and Will.” In *The Blackwell Guide to Descartes Meditation*, edited by Stephen Gaukroger, 142–159. Malden, MA: Blackwell Press.
- Della Rocca, Michael. 2011. “Taking the Fourth: Steps toward a New (Old) Reading of Descartes.” *Midwest Studies In Philosophy* 35: 93–110.
- Kenny, Anthony. 1972. “Descartes on the Will.” In *Cartesian Studies*, edited by R. J. Butler, 1–31. Oxford: Blackwell.
- Naaman-Zauderer, Noa. 2010. *Descartes' Deontological Turn*. New York: Cambridge University Press.
- Newman, Lex. 1999. “The Fourth Meditation.” *Philosophy and Phenomenological Research* 59 (3): 559–591.