

## WHAT DOES THE PREMISE “*A DECEIVER DECEIVES ME*” CONCLUDE? – DESCARTES’ DECEIVER ARGUMENT RECONSIDERED<sup>1</sup>

AYUMU TAMURA, University of Tsukuba, Ibaraki, Japan

TAMURA, A.: What does the Premise “a Deceiver Deceives me” Conclude? –  
Descartes’ Deceiver Argument Reconsidered  
FILOZOFIA, 74, 2019, No 4, pp. 308 – 317

Descartes insists, “[...] there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that case I too undoubtedly exist, if he is deceiving me [...]” (AT-VII, 25; CSM-II, 17). In what way can we draw evidence that our existence can be drawn from our being deceived? The interpretations that the earlier studies have shown is not a monolith. Then I will search for some inherent characteristics of deception, and analyse the construction of the reasoning, “if a deceiver deceives me, I exist.” To be concrete, I show: 1) that it is not “I exist” but “I think” that is concluded from the supposition “a deceiver deceives me/I am deceived by a deceiver”; 2) that the attributions of the self as a thinking thing, i.e. understanding, affirming, and denying (AT-VII, 28; CSM-II, 19), are discovered in the very supposition of a deceiver.

**Keywords:** Descartes – Meditationes de prima philosophia – Cogito – Deceiver argument – Skepticism

### 1. Introduction

Science holds true under the condition that the human intellect is genuine (in the sense that it can grasp truths). Science would lose all the significance and validity if the human intellect had not been created with the ability to grasp them. According to the ancient Skeptic, especially Pyrrhonism, we cannot make a firm judgement on anything and therefore cannot construct science. Descartes makes use of the views of the Skeptic to lay the foundation of his system of metaphysics in the First and Second Meditations, as it is well-known. The following is where Descartes’ methodological doubt is deepened extremely, and the most crucial for this study:

---

<sup>1</sup> I wish to thank Dr. Yoshinori Tsuzaki, Associate Professor at University of Tsukuba, for helpful discussions. Readers can find more on this topic in another article of mine that was published in Japanese: 田村, 歩. (2018): デカルトにおける〈欺かれる私〉について——欺かれるという事態からは何が帰結するのか. 哲學, 69, 200-214.

**[I]** But I have convinced myself that there is absolutely nothing in the world, no sky, no earth, no minds, no bodies. Does it now follow that I too do not exist? No: if I convinced myself of something then I certainly existed. **[II]** But there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that case I too undoubtedly exist, if he is deceiving me; and **[III]** let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. (AT-VII, 25; CSM-II, 16-7)<sup>2</sup>

In this, by convincing himself that there is no sky, no earth, no minds, no bodies (**I**), and by supposing that a deceiver of supreme power and cunning deceives himself (**II**), Descartes proves that he exists on the condition that he thinks (**III**). In this paper, I focus on the argument from the supposition of a deceiver. According to Gouhier, the deceiver is “a kind of methodological puppet,” “a methodological artifice to detain us in the doubt until we encounter with a first truth without falling into a trap of the probable.” The first truth is “I am, I exist,” because however powerful the deceiver is, for him to deceive us, we must exist (Gouhier 1973, 155-157). Then, in what way can we draw or find evidence that our existence can be drawn from our being deceived by a deceiver? In Gouhier’s view, it is not clear. At a glance, it seems that passage **II** says that we exist *as an object of deception* because for a deceiver to deceive someone, he, who is deceived, must exist. Some commentators interpret it in this way: Kenny, for example, insists that the supposition of a deceiver is not related to the self as a subject of thinking, but to the self as an object of deception, and that the self as such [latin: *me*] must represent something (Kenny 1987, 57). Similarly, Pariente insists that if X does not exist, X is not the victim of deception; if X is the victim of deception, X exists, and draws the existence of the self based on the relation between an actor (a deceiving thing) and an actee (a deceived thing) (Pariente 1999, 35).

On the other hand, Dicker insists, “while it would be possible for a powerful deceiver to fool Descartes into thinking that he had a body even if he did not have one, it would be impossible for any deceiver, no matter how powerful, to fool Descartes into thinking that he was thinking when he was not thinking - for thinking that one is thinking is thinking” (Dicker 2013, 82). Moreover, Wagner insists that the “I” exists, not as a passive thing of an object of deception, but as a positive thing that exercises the causal power in generating the idea of a deceiver and the ideas of all the other things. According to him, all ideas in our minds are generated by means of the causal power of our minds, and the idea of a deceiver is just one of them (Wagner 2014, 110 – 112).

---

<sup>2</sup> In accordance with current practice, I use the abbreviation “AT” for the standard edition of Descartes’ works (Descartes 1964 – 1974), and “CSM” for the standard translation (Descartes 1985).

As above, the interpretations that the earlier studies have shown is not a monolith. What did Descartes intend to conclude on the supposition “a deceiver deceives me”? In this paper, I will try to find out what is concluded from being deceived; it is not until we find out that we can finally understand Descartes’ thought on the *cogito*. Then I will search for some inherent characteristics of *deception*, and analyse the construction of the reasoning, “if a deceiver deceives me, I exist.” To be concrete, I show: 1) that it is not “I exist” but “I think” that is concluded from the supposition “a deceiver deceives me/I am deceived by a deceiver”<sup>3</sup>; 2) that the attributions of the self as a thinking thing, i.e. *understanding*, *affirming*, and *denying* (AT-VII, 28; CSM-II, 19), are discovered in the deceiver argument.

This paper has a narrow scope and specialized object, but it shall provide the way to interpret the deceiver argument in the Second Meditation more precisely than ever before.

## 2. Examination of the traditional interpretation

In this section, I examine the validity of the interpretation that the “I” in the passage II exists as an *object of deception*. Here, I divide the argument of the Second Meditation, which I quoted at the beginning of this paper, into three parts:

- (1) If I convince myself that there is nothing, I exist.
- (2) If a deceiver deceives me, I exist.
- (3) So long as I think, I exist.

The first argument corresponds to “if I convinced myself of something then I certainly existed,”<sup>4</sup> the second corresponds to “I too undoubtedly exist, if [a deceiver] is deceiving me,” and the third corresponds to “[the deceiver] will never bring it about that I am nothing so long as I think that I am something”; such is the structure of the argument of the *cogito* in the Second Meditation. The questions here: what kind of relation is there between the first and second? What kind of relation is there between the second and third? Can we consider the “I” in the second as a mere object of deception or not?

Firstly, what kind of relation is there between the first and second? One likely explanation is that they are separate and unconnected to each other. In this case, if we convince ourselves, we exist because for us to convince ourselves, we must exist; if

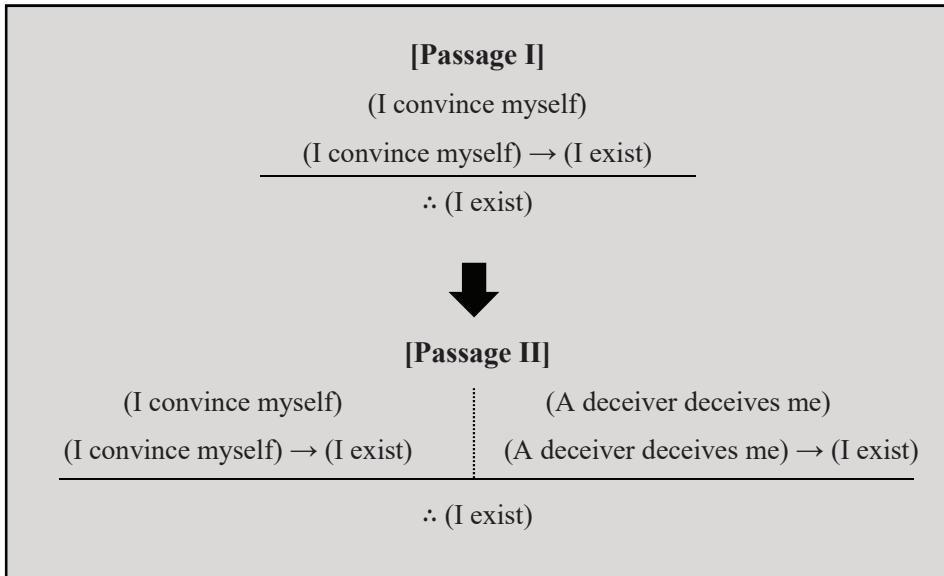
---

<sup>3</sup> Indeed this conclusion is similar to Dicker’s one that the deceiver hypothesis cannot disprove the certainty of “I think,” but I will deal with the reason that this hypothesis can draw the certainty of “I think” in more detail.

<sup>4</sup> The original is past-tense; but I will not make distinction of tense.

a deceiver deceives us, we exist because for him to deceive us, we must exist. In this way, these two arguments draw the same conclusion of “I exist” having no influence on each other. Figure 1 shows this relation:

Fig. 1



However, it is impossible to consider the first and second to be separate and unconnected to each other. *In general*, it seems that “I exist” can be established in the passage I, because “I convince myself” is one of the modes of “I think,” and “I think” is the condition of “I exist.” *In the context of Descartes*, however, it is not until we overcome a menace of a deceiver that we can draw our existence from “I convince myself” and “I think.” Descartes writes, “if I convinced myself of something then I certainly existed,” and adds, “*but* there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me.” Here, we pay attention to a junction “*but [sed]*.” This junction signifies that even the conviction that we can convince ourselves of something is as subjected to the methodological doubt, as the conviction that we have a body, or that we can sense something, and so on. That is, the supposition of a deceiver of supreme power provides the possibility that although we do not convince ourselves, we are deceived by a deceiver into believing that we convince ourselves of something. Moreover, Descartes writes, “[...] I myself may perhaps be the author of these thoughts? In that case am not I, at least, something?” It is true that being the author of some thoughts is thinking of something, so at first

glance, the existence of the self seems to be drawn from being “the author of these thoughts”; however, Descartes denies this way and convinces himself that there is nothing. Therefore, both thinking of something (i.e. being the author of some thoughts) and convincing oneself of something cannot assure the certainty of one’s existence until a release of the supposition of a deceiver.

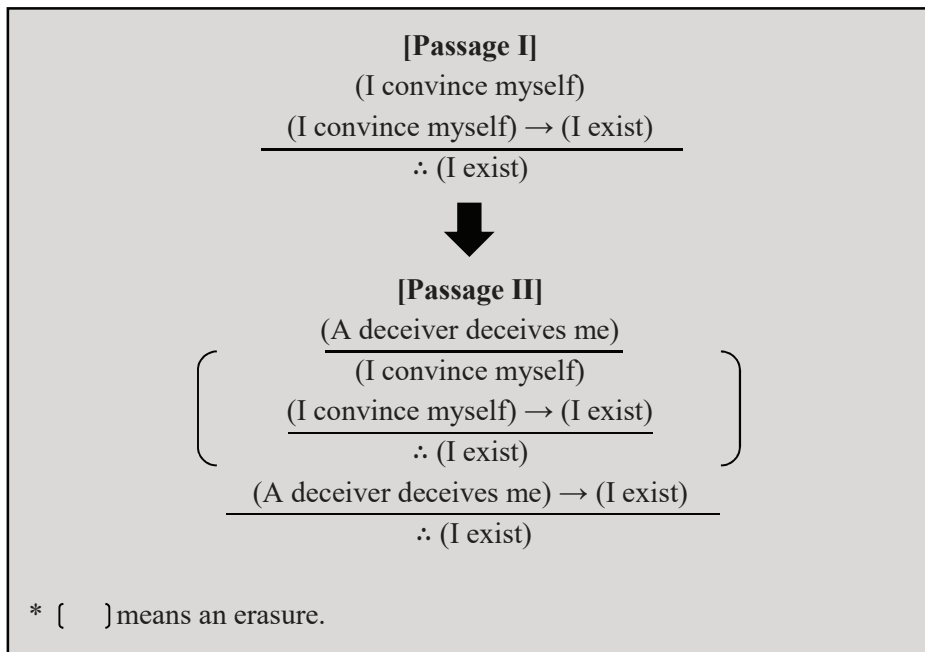
As stated above, the status of one’s own mental acts is not privileged. This way of thinking owes its origin to Pyrrhonism. Montaigne writes in his *Essays*, referring to the *Outlines of Pyrrhonism* of Sextus (The *Essays* was Descartes’ favorite book):

They only put out their propositions to contend with those they think we have in our belief. If you take their arguments, they will as readily maintain the contrary; ’tis all one to them; they have no choice. If you maintain that snow is black, they will argue, on the contrary, that it is white; if you say it is neither the one nor the other, they will maintain that ’tis both. If you hold, as of certain judgment, that you know nothing of it, they will maintain that you do: Yes, and if, by an affirmative axiom, you assure them that you doubt, they will argue against you that you doubt not, or that you cannot judge and determine that you doubt. And by this extremity of doubt, which jostles itself, they separate and divide themselves from many opinions, even of those that have several ways maintained doubt and ignorance (Montaigne 1930, 196).

According to Montaigne, skeptics can have refutations to all judgements; they can refute that we doubt. Thus they should be able to doubt that we convince ourselves. But, for what reason? Descartes always needs firm *reasons to doubt* for the methodological doubt, which “is not [based on] a flippant or ill-considered conclusion, but is based on powerful and well thought-out reasons” (AT-VII, 21; CSM-II, 15). To doubt because Montaigne (or skeptics) tells to do so is to doubt based on *thoughtless* reasons; to avoid doing so, Descartes uses the supposition “a deceiver deceives me” as the reason to doubt “I convince myself” and “I am the author of some thoughts.”

From this point of view, to accept the supposition of a deceiver is to reject the validity of all knowledge, including the reasoning “if I convince[d] myself, I exist[ed]” and “if I am the author of some thoughts, I am something”; thus, the first and second arguments are not separate and unconnected to each other. Figure 2 shows this relation:

Fig. 2



As Fig. 2 shows, the premise “a deceiver deceives me” must be held to assure the certainty of the conclusion “I exist” because this premise makes the first argument *invalid*. The superiority of the second over the first is kept until God’s honesty is proven. Since Descartes keeps the supposition of a deceiver through the whole of the Second Meditation, the premise “I convince myself” cannot draw the certainty of the existence of the self. In passage II, only the premise “a deceiver deceives me” can draw it<sup>5</sup>.

When we interpret in this way, can we consider the self as an object of deception? No: if we considered it as such, it would be very difficult for us to make the way from “if a deceiver deceives me, I exist”(2) to “so long as I think, I exist”(3), and the way from “I exist” to “I am a thinking thing.” For Descartes, “I exist” is simply not sufficient; it is more important that the “I” exists *as a substance separated from a body, whose nature is only thinking*, as Descartes says, “I am a thinking thing” (AT-VII, 34) and “I think, therefore I am” (AT-VI, 32 – 33; AT-VII, 140; AT-VIII, 7 – 8; AT-X, 523). In passages II and III, Descartes writes, “I too undoubtedly exist, if [a deceiver] is deceiving me,” and adds, “he will never bring it about that I am nothing *so long as I think* that I am something.” In this, the condition “so long as I think [...]”

<sup>5</sup> For this reason, we cannot accept Kenny’s insistence that the premise “a deceiver deceives me” is *hypothetical*.

plays a necessary part of preparation for the advance from “I exist” to “I am a thinking thing.” However, if we considered the “I” in passage II as an object of deception, we would miss the relation between “if a deceiver deceives me” and “so long as I think,” and would not advance from the second argument to the third one. In the next section, I will analyse the relation between “if a deceiver deceives me” and “so long as I think.”

### 3. Thought as an object of deception

How can we insist that we exist *so long as we think*, based on the previous reasoning that we exist *if we are deceived*? To solve this question, I will focus on some inherent characteristics of *deception* in this section. It is true that Descartes’ deceiver/deceiving God argument is based on the traditional arguments<sup>6</sup>, of course; however, it seems that Descartes was aware of some inherent characteristics of *deception*, which his predecessors had not.

\*

What kind of act is *deception*? It is defined as, to make someone believe a falsehood to be a truth, or to make someone believe a truth to be a falsehood. Then what is a requirement for being deceived? – *understanding*, *affirming*, and *denying*. When we are deceived into believing that A is B although A is actually not B, we must affirm that A is B and deny that A is NOT B; moreover, we must understand the meaning of the proposition “A is B.” To be concrete, when we are deceived into believing that the earth is flat, for example, we must understand the meaning of “the earth is flat,” and affirm that the earth is flat, or deny the contrary (“the earth is NOT flat”). Thus, it is not until we *understand*, *affirm*, and *deny* something that we can be deceived<sup>7</sup>. In other words, our being deceived necessarily accompanies our understanding, affirming, and denying, i.e. our *thinking*<sup>8</sup>. Conversely, it is impossible to deceive things that do not think, i.e. matter. A deceiver schemes to alter the content of our thought, and the more he succeeds, the more certainly we think.

From the point of view above, the premise “a deceiver deceives me/I am deceived by a deceiver” affirms “I think (since no one can be deceived without *understanding*, *affirming*, and *denying*, i.e. *thinking*),” not “I exist (since for a deceiver to deceive

---

<sup>6</sup> Cicero and Suárez, who influenced Descartes, dealt with the deceiving God. See Brown (2013), 25 – 42; Menn (2002), chap. 5, sec. B; Secada (2000), 44.

<sup>7</sup> Even though we try to deceive a young child of this matter, he/she cannot be deceived if he/she does not understand the meanings of the *earth* and *flat*.

<sup>8</sup> These three are modes of thought, as Descartes writes, “but what then am I? A thing that thinks. What is that? A thing that doubts, *understands*, *affirms*, *denies*, is willing, is unwilling” (AT-VII, 28; CSM-II, 19). It is remarkable that these modes of thought, which are presented in the middle of the Second Meditation, are discovered in the deceiver argument.

someone, he/she must exist as an object of deception).” It is clear from his declaration, “let [the deceiver] deceive me as much as he can” that Descartes thought in this way: even after Descartes succeeded in drawing the certainty of his existence, he did not say “the deceiver does not always deceive me,” or “the deceiver failed to deceive me,” but “let him deceive me as much as he can”; this is his expression that we need not care of the deceiver any more and challenge to him. [Descartes leaves the deceiver as it is, although Gouhier and Wagner turn it into a product of will or imagination.] The deceiver can always deceive and mislead us, but every time he does so, he gives us the evidence that we think.

In the end, the deceiver argument is a tool to draw the certainty of “I think.” In this reading, we can understand the relation between “if a deceiver deceives me” and “so long as I think” clearly, and advance from the second argument to the third one. We cannot reach “I exist” at a stroke from “if a deceiver deceives me”; we first reach “I think” from this premise, and if we think, we exist because we experience in ourselves that we cannot think unless we exist<sup>9</sup>. Then we cannot support Alquié’s opinion anymore that in the *Meditations*, unlike in the *Discourse on Method* and the *Principles of Philosophy*, Descartes puts existence above thinking (Alquié 1956, 92). In fact, in the *Meditations*, Descartes explains the indubitability of “I think,” from which the certainty of “I exist” is drawn, even more minutely than in all of the other works. In the Second Meditation, we must grasp this two-step order of thought: (1) from “I am deceived” to “I think,” and (2) from “I think” to “I exist.” The transition from “I am deceived” to “I think” is possible only after deliberation on the state of mind at the moment to be deceived. Through this deliberation, Descartes grasped the inherent characteristics of deception precisely, and could show the certainty of the *cogito* without excluding the deceiver: “let him deceive me as much as he can.” In this respect, Descartes is in contrast to Francisco Suárez, who denies instantly the possibility that God deceives us because God cannot deceive *by its very nature*, without deliberation on the state of mind at the moment to be deceived (Suárez 1614, IX, sec 2, vii).

#### 4. Conclusion

The following are the conclusions in this paper:

---

<sup>9</sup> Descartes writes, “when someone says ‘I am thinking, therefore I am, or I exist’, he does not deduce existence from thought by means of a syllogism, but recognizes it as something self-evident by a simple intuition of the mind. This is clear from the fact that if he were deducing it by means of a syllogism, he would have to have had previous knowledge of the major premiss ‘Everything which thinks is, or exists’; yet in fact he learns it from experiencing in his own case that it is impossible that he should think without existing” (AT-VII, 140; CSM-II, 100).



- 1) The first argument (“if I convince myself of something, I exist”) and the second argument (“if a deceiver deceives me, I exist”) are not separate and unconnected; the premise of the latter argument makes the former argument invalid.
- 2) The “I” drawn from the second argument cannot be an object of deception because if so, we could not understand the relation between the second argument (“if a deceiver deceives me, I exist [*as an object of deception*]”) and the third one (“*so long as I think*, I exist”).
- 3) The conclusion we can draw from the second argument is not “I exist” but “I think”; furthermore, we can discover in passage II the attributions of the self as a thinking thing, *understanding*, *affirming*, and *denying*, which are presented in the middle of the Second Meditation.

Lastly, I comment on the *cogito* as the *first principle*. Descartes considers the *cogito* to be “so firm and sure that all the most extravagant suppositions of the sceptics” (AT-VI, 32; CSM-I, 127) and “the most certain and evident of all” (AT-VII, 25; CSM-II, 17). It is not only because the *cogito* can be discovered at first that it can be the first principle; it is also because only the *cogito* can be established *even if we are deceived by a deceiver of supreme power*. The certainty of all the other things, i.e. the existence of matter, mathematical truths, etc. is not possible until God’s honesty is proven. However, we are certain of “I think” every time we are afraid that a deceiver deceives us. Descartes repeats, “let him deceive me as much as he can,” and shows the certainty of the *cogito without excluding the deceiver*; The fact that we can establish the *cogito* without excluding him is the very proof of the *cogito* being the first principle.

## Bibliography

- ALQUIÉ, F. (1956): *Descartes: l’homme et l’œuvres*. Paris: Hatier-Boivin.
- BROWN, D. (2013): Descartes and Content Skepticism. In: *Descartes’ Meditations: A Critical Guide*. Karen Detlefsen (ed.). Cambridge: Cambridge University Press, 25 – 42.
- DESCARTES, R. (1964 – 1974): *Œuvres de Descartes*. 13 vols. Adam. Ch. – Tannery, P. (eds.). Paris: J. Vrin. [AT]
- DESCARTES, R. (1985): *The Philosophical Writings of Descartes*. 2 vols. Transl. Cottingham, J., Stoothoff, R., Murdoch, D. Cambridge: Cambridge University Press. [CSM]
- DICKER, G. (2013): *Descartes: An Analytic and Historical Introduction*. New York – Oxford: Oxford University Press.
- GOUHIER, H. (1973): *Descartes: Essais sur le «Discours de la méthode», la métaphysique et la morale*. Paris: J. Vrin.
- KENNY, A. (1987): *Descartes: A Study of His Philosophy*. New York: Random House.
- MENN, S. (2002): *Descartes and Augustine*. Cambridge: Cambridge University Press.

- MONTAIGNE, M. (1930): *The Essays of Michel de Montaigne*. Vol. II. Transl. Cotton, Ch. London: G. Bell.
- PARIENTE, J.-C. (1999): La première personne et sa fonction dans le *Cogito*. In: *Descartes et la question du sujet*. Kim Sang Ong-Van-Cung (ed.). Paris: P. U. F., 11 – 48.
- SECADA, J. (2000): *Cartesian Metaphysics: The Scholastic Origins of Modern Philosophy*. Cambridge: Cambridge University Press.
- SUÁREZ, F. (1614): *Metaphysicarum disputationum*. Moguntae: Mylius Bireckmannus.
- WAGNER, S. I. (2014): *Squaring the Circle in Descartes' Meditations: The Strong Validation of Reason*. Cambridge: Cambridge University Press.

---

Ayumu Tamura  
University of Tsukuba  
1-1-1 Tennodai, Ibaraki  
Japan  
e-mail: des.ayumu.tamura.cartes@gmail.com