THOUGHT EXPERIMENT ANALYSES OF RENÉ DESCARTES’ COGITO¹

C. P. Hertogh²

ABSTRACT: René Descartes’ Cogito is an example of a paradigmatic thought experiment, herald of both subjectivism and new science in Europe’s Modern Age, that seems to have escaped the attention of thought experiment philosophers. On deep analysis, the Cogito appears as universal instantiation (or modus ponens with implicit major ‘whatever has the property of thinking, exists’). The Cogito has strong rhetorical effects for it narratively generalizes from I to all human kind, and its historical and philosophical success can be explained from its concise enthymematic structure that rings true in many possible senses. We consider it a preeminent example of a thought experiment as it states the power of thinking as its very contents. From Descartes’ methodology of doubt we can conclude that, e.g., on a Wittgensteinian interpretation, the Cogito is a logical thought experiment rather than a psychological one.


1 DESCARTES’ DOUBT THOUGHT EXPERIMENT

The famous thought experiment of René Descartes (1596-1650) appears in the beginning of part IV of Discourse de la méthode. It is also known as Descartes’ Doubt Thought Experiment, and we quote it here in full from the Gutenberg online English translation by John Veitch (DESCARTES, 1901):

[…] as I then desired to give my attention solely to the search after truth, […] and that I ought to reject as absolutely false all opinions in regard to which I could suppose the least ground for doubt, in order to ascertain whether after that there remained aught in my belief that was wholly indubitable. Accordingly, seeing that our senses sometimes deceive us, I was willing to suppose that there existed nothing really such as they presented to us [TE]; and because some men err in reasoning, and fall into paralogisms, even on the simplest matters of geometry, I, convinced that I was as open to error as any other, rejected as false all the reasonings I had hitherto taken

¹ This article on René Descartes’ Cogito thought experiment is derived from Ch. 6 of Semantics of Thought Experiments (HERTOGH, 2015, p. 199-206). Please, see Acknowledgements.

http://dx.doi.org/10.1590/S0101-31732016000300002

² C. P. Hertogh (e-mail noromyxo2005@gmail.com) studied Philosophy at VU University (BPhil), UvA University, Amsterdam (MPhil), and VUB University, Brussels (PhD Thesis Semantics of Thought Experiments—under review).
for demonstrations \( [\text{TE}_1] \); and finally, when I considered that the very same thoughts (presentations) which we experience when awake may also be experienced when we are asleep, while there is at that time not one of them true, I supposed that all the objects (presentations) that had ever entered into my mind when awake, had in them no more truth than the illusions of my dreams. \( [\text{TE}_2] \) But immediately upon this I observed that, whilst I thus wished to think that all was false, it was absolutely necessary that I, who thus thought, should be somewhat; and as I observed that this truth, I think, therefore I am (COGITO ERGO SUM), was so certain and of such evidence that no ground of doubt, however extravagant, could be alleged by the sceptics capable of shaking it, I concluded that I might, without scruple, accept it as the first principle of the philosophy of which I was in search. \( [\text{TE}_3] \). \( \text{(DESCARTES, 1901, Pt. IV--bracketed counting of sub-thought experiments added, \( [\text{TE}_1] \) … \( [\text{TE}_4] \))} \).

Its core line ‘I think, therefore I am’, first appeared in French, ‘je pense, donc je suis’ in the 1637 edition of Discourse de la méthode. The Latin translation ‘cogito ergo sum’ was not parenthetically added to the first 1637 edition, but it first appeared in Descartes’ 1644 Principia Philosophiae (Principles of First Philosophy).\(^3\) \( \text{TE} \) indicators, which mark the different sub-thought experiments, are, e.g., ‘to which I could suppose’ (‘en quoi je pourrais imaginer’), ‘I was willing to suppose’ (‘je voulus supposer’), ‘I, convinced that I was’ (‘jugeant que j’étais’), ‘when I considered that’ (‘considérant que’), ‘whilst I thus wished to think’ (‘pendant que je voulais ainsi penser’).

\(^3\) The French original text from the first 1637 edition (DESCARTES, 1637) reads like this: […] pour ce qu’alors je désirais vaquer seulement à la recherche de la vérité, […] et que je rejetasse, comme absolument faux, tout ce en quoi je pourrais imaginer le moindre doute afin de voir s’il ne resterait point, après cela, quelque chose en ma créance, qui fût entièrement indubitable. Ainsi, à cause que nos sens nous trompent quelquefois, je voulus supposer qu’il n’y avait aucune chose qui fût telle qu’ils nous la font imaginer. Et pour ce qu’il y a des hommes qui se méprennent en raisonnant, même touchant les plus simples matières de géométrie, et y font des paralogismes, jugeant que j’étais sujet à faillir, autant qu’aucun autre, je rejetai comme fausses toutes les raisons que j’avais prises auparavant pour démonstrations. Et enfin, considérant que toutes les mêmes pensées, que nous avons étant éveillés, nous peuvent aussi venir, quand nous dormons, sans qu’il y en ait aucune, pour lors, qui soit vraie, je me résolus de feindre que toutes les choses qui m’étaient jamais entrées en l’esprit n’étaient non plus vraies que les illusions de mes songes. Mais, aussitôt après, je pris garde que, pendant que je voulais ainsi penser que tout était faux, il fallait nécessairement que moi, qui le pensais, fusses quelque chose. Et remarquant que cette vérité: je pense, donc je suis, était si ferme et si assurée, que toutes les plus extravagantes suppositions des sceptiques n’étaient pas capables de l’ébranler, je jugeai que je pouvais la recevoir, sans scrupule, pour le premier principe de la philosophie que je cherchais. (DESCARTES, 1637, p. 32-33).

The technical term ‘thought experiment(s)’ is abbreviated to \( \text{TE} \), and appears in some compounds as \( \text{TE} \) Matrix, \( \text{TE} \) premise, \( \text{TE} \) inference, \( \text{Doubt TE} \).
In fact, Descartes’ *Doubt TE* is a multiple thought experiment, consisting of a series of, e.g., four sub-thought experiments, to which *Cogito ergo sum* is the conclusion.

### 2 Logical Analysis of Deep Structure: Modus Ponens with Implicit Major

The *Cogito* is an enthymeme (or unfinished syllogism) and its analyses are simpler than general formulas of thought experiment analyses may suggest, as it can be explained as modus ponens with a hidden major. On logical analysis the core line appears as minor of the argument, since it is an individual rather than a general statement: “I think, therefore I am”.

Applying the methodology of TE Matrix (Thought Experiment Matrix) as developed in our PhD dissertation, we first bracket the formalization of the enthymatic thought experiment to emphasize that it is not a logically valid argument yet. Applying predicate logic, bracketing the so-called TE premise or TE minor:

#### [1]

**Suppose**

<table>
<thead>
<tr>
<th>Tx</th>
<th>x Thinks, x is Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>constant, e.g. Descartes, I, etc.</td>
</tr>
<tr>
<td>Ex</td>
<td>x Exists/Is</td>
</tr>
</tbody>
</table>

(1) 

\[
\begin{array}{c|c|c}
& Ta & \text{a thinks, e.g. I think, Descartes thinks} \\
\hline
& ------ & \text{TE inference} \\
\hline
\end{array}
\]

(2) 

\[
\begin{array}{c|c|c}
& Ea & \text{a exists, e.g. I exist/am, Descartes exists/is} \\
\hline
\end{array}
\]

According to our semantic extended argument view of thought experiments, thought experiments are incomplete arguments (like Descartes’ enthymeme), that can be completed by exemplification of tacit, contextual or theoretical premises and presuppositions (axioms, theorems, derivations in mathematics; rules, regularities, epistemological and scientific principles in empirical sciences). In possible worlds terminology, when possible worlds can be defined as constituents of theories, the thought experiments need to pick out available and accessible possible worlds (as from axioms to principles) that suffice to complete and validate the argument unto a formally and informally
logically valid and sound argument (e.g., inference is valid, premises and conclusion are true—argument is sound).

In Descartes’ *Cogito* it is quite easy to find the major principle, as it is in the same passage of the *Doubt TE*: “[...] I concluded that I might, without scruple, accept it as the first principle of the philosophy of which I was in search.” (DESCARTES, 1901).

The ‘principle’ can be summarized like this (CIPROTTI, 2009, p. 65):

Whatever/Whoever thinks, exists.

We can develop the full argument by adding the major to the enthymematic thought experiment--

\[
\begin{align*}
\text{Suppose} & \\
\forall x & \text{universal quantifier} \\
\Rightarrow & \text{if } \ldots, \text{then } \ldots \\
(0') & \forall x (Tx \rightarrow Ex) \text{ whatever thinks, exists (for all } x \text{: if } x \text{ thinks, then } x \text{ exists)} \\
(1') & Ta \text{ a thinks, e.g., I think, Descartes thinks} \\
(2') & Ea \text{ a exists, e.g., I exist/am, Descartes exists/is}
\end{align*}
\]

From narratology we know that the narrator does not need to be identical to the author, and from the context of Descartes’ theory and the history of philosophy we know that *a* has been interpreted to refer to any human (or sentient) being. For these reasons the TE can be expressed with help of universal quantification in the major and existential quantification in both the minor and conclusion:

\[
\begin{align*}
\text{Suppose} & \\
\exists x & \text{existential quantifier} \\
(0'') & \forall x (Tx \rightarrow Ex) \text{ whoever thinks, exists} \\
(1'') & \exists x Tx \text{ there is/are at least one/some human being(s) that think(s)} \\
(2'') & \exists x Ex \text{ there is/are at least one/some human being(s) that exist(s)}
\end{align*}
\]
Formula [1] is the ‘raw’ thought experiment in terms of Sorensen (1992, p. 5). We have bracketed it because it is not a logically valid or sound argument yet. Formulas [2′] and [2″], however, are valid argument schemes, based on modus ponens; [2′] is universal instantiation on predicate logical analysis, which simply shows an example of a universal rule, regularity, law etc. Note that the thought experiment is easier to remember than the full argument, and it catches the attention by rhetorical use of I, in fact, signifying everyone, every human, etc.

The thought experiment operator or inference, may be equivalent to the psychological faculty of intuition according to many thought experiment philosophers such as Ernst Mach, Thomas Kuhn, Daniel Dennett, Tamar Gendler, John Norton, Roy Sorensen. The logical effect of the Cogito, in particular, is equivalent to modus ponens, universal instantiation, or material implication (elimination). Descartes held that ‘I think, therefore I am’ is self-evident, implying that is it like a (necessary) truth, so modus ponens and mathematical logical operations are appropriate. In addition, upon analysis the full Cogito argument appears self-evident because it is like petitio principii (begging the question, cf. some analyses of the Anselmian Ontological Argument); this, of course, is not a formal fallacy, for the inference is logically valid. Furthermore, as rhetorical analyses may show, Descartes is, in fact, stating a universal law or rule by a compelling and striking instantiation which is exactly what universal instantiation is meant to do by definition.

3 Surface Structure and Rhetorical Analysis

The Doubt TE can be considered an anti-skeptical argument because, e.g., ‘the sceptics’ are mentioned towards the end of the thought experiment

4 Modus ponens, \( p \rightarrow q, p, \) so, \( q \), belongs to the propositional calculus. Existence as a predicate (e.g. \( x \) Exists/Is) is not excluded in, e.g., Kripke logic, and it can also be applied to analyses of Anselm’s Ontological Argument. On predicate logical analysis, the full Cogito argument is universal instantiation – e.g., as applied by Rudolf Carnap (1966, p. 7, 17) in a generic scheme of scientific explanation or prediction, which fits in with modus ponens (instantiation):

1. \( (x)(Px \rightarrow Qx) \)
2. \( Pa \)
3. \( Qa \)

Note that instead of deductive inference, the Cogito and many scientific rules, regularities etc. could also be stated as induction, e.g., from the experience that he exists from the very fact of his doubting or thinking, Descartes may have induced to whoever thinks, exists, e.g. \( Ta \rightarrow Ea, \) so, \( \forall x (Tx \rightarrow Ex) \).

Cf. Carnap’s scheme reads both as explanation and prediction.
text. Introducing the final core argument, ‘I think, therefore I am’ (TE₁), there
are three preceding sub-thought experiments, wherein Descartes is granting
the skeptics quite a lot: “[…] that I ought to reject as absolutely false all
opinions in regard to which I could suppose the least ground for doubt”. (DESCARTES, 1901).

The first sub-thought experiment (TE₁) concludes to the epistemological
failure of sense experience, ‘that there existed nothing really such as they [our
senses] presented to us,’ from the very fact that they ‘sometimes deceive us.’
The second sub-thought experiment (TE₂) concludes to the epistemological
failure of reason and mathematics, “I […] rejected as false all the reasonings
I had hitherto taken for demonstrations,’ from the very fact that ‘some men
err in reasoning, and fall into paralogisms, even on the simplest matters of
geometry.” (DESCARTES, 1901). The third sub-thought experiment (TE₃)
concludes to the epistemological failure of all content of the human mind, ‘all
the objects (presentations) that had ever entered into my mind,’ from the very
fact that ‘the very same thoughts (presentations) which we experience when
awake may also be experienced when we are asleep, while there is at that time
not one of them true.”

Rhetorical analyses may reveal multi-interpretability of the ‘je pense,
donc je suis.’ Deep analyses show that Descartes wants to say (as a general
rule, law) that whatever has the property of thinking exists merely because
of thinking. However, Descartes says ‘I think, therefore I am’ (instead of
‘whatever thinks, exists’) for a manifold of (rhetorical) reasons as rewarded by
history afterwards.

1. Part of the rhetorical force of Descartes’ thought experiment comes from
the use of I, changing a universal law into a seeming subjective statement.
This rhetorical use of I is well-known and we may use it in daily conversation,
too. E.g., if A wants to persuade B to do T, A could say:

I have also done T.
or
If I were you, I would do T.

Footnotes:
[5] Dreams can constitute thought experiments, too. E.g. Zhuangzi’s Butterfly Dream, Hundred Schools
of Thought Period, c. 369-286 BCE (see e.g. ZHUANGZI, 2012, HERTOGH, 2013), Martin Luther King Jr.’s I Have a Dream, 1963 (see e.g. KING, 1963, HERTOGH, 2015, Section 1.4, p. 52-
59), Albert Einstein’s daydreaming e.g. Chasing a Beam of Light, 1949 etc. (EINSTEIN, 1969/1949,
2. Many accounts of the history of philosophy have credited this phrase as the beginning of subjectivism in philosophy, that is, subjectivism as a universal theory.

3. At the same time, the phrase is considered a concise expression of mind-body dualism saying that thinking, i.e., mind (res cogitans), is indispensable for existing, i.e., more essential than the apparent body (res extensa) that can be verified by sense experience.

4. We may suppose that Descartes did not want to say whatever is thinking, but whoever is thinking – possibly including animals and plants too (all of the organic world) – without leaping into pantheism (excluding the inorganic world), though by the use of the prominent I he may have wanted to restrict it to humans particularly. Nowadays, in our digital global village and global cross-cultural societies, it embraces all races, ethnicities, genders, cultures, religions, ages, abilities, sexual orientations, births, social standings, etc.

5. On a predicate logical interpretation of universal instantiation, the thought experiment argument has the logical strength of a scientific explanation or prediction (cf. note 4). It links the paradigmatic phrase ‘I think, therefore I am’ to the scientific methodology of the European Modern Age.

6. The universal law ‘whoever has the property of thinking, exists’ could – in another sense of to exist (être, esse) – relate to the difference between life and death, saying that whoever is thinking – i.e., has a functional mind or brain – is alive, and whoever is not thinking anymore, whose mind or brain is not functioning anymore, is dead. When relating it to the difference between life and death, the statement ‘whoever is thinking, exists’, as opposed to ‘whatever is not thinking anymore, is dead’, is like a universal physical or biological law, not just a probabilistic inference (as Aristotle categorized enthymemes as rhetorical syllogisms), but indeed a scientific deductive syllogism about necessary propositions. Although it is

---

*In his contribution to Sharif 1963, *A History of Muslim Philosophy*, Fazlur Rahman, who edited the Arabic text of Ibn Sinâ’s *De Anima* (Rahman 1959), quotes from Descartes *Meditations* II to show that the existence of the self in Descartes’ *Cogito* is like a logical necessity from inference – a near equivalence of ‘I think’ and ‘I am’ (res cogitans) – rather than a contingent fact as an immediate datum of consciousness, as it is in Ibn Sinâ’s *Flying Man*: What of thinking? I find here that thought is an attribute that belongs to me; it alone cannot be separated from me. I am, I exist, that is certain. But how often? Just when I think; for it might possibly be the case, that if I ceased entirely to think, I
not in accordance with traditional interpretations of Descartes, we cannot confirm that the immortality of the soul can be immediately derived from the *Cogito* thought experiment. Even the subtitle of the *Meditationes* was changed from demonstrating the immortality of the soul (‘[…] *animae immortalitas demonstrantur*’ in first edition of 1641) to demonstrating only the distinction of the soul from the body (‘[…] *animae humanae à corpore distinctio demonstrantur*’ in the second edition of 1642).

7. It is a thought experiment that says that thinking, performing thought experiments like ‘*je pense, donc je suis*’, amounts to as much as to be, to exist, to be alive, to be conscious, to be human – and by *I* to be individual, subjective, and so on, stating a paradigmatic maxim that is still a leading one to many citizens of twenty-first century societies.

8. The ‘magic’ comes from its many senses condensed to only one phrase as short as a formula. The concise and incomplete, but compelling and striking form turns ‘*je pense, donc je suis*’ into a preeminent example of a thought experiment, even stating the power of thinking by its very content. It is an example of a strong thought experiment that can be easily explicated by its context and may serve as an adage as well.

9. The thought experiment interpretation of the *Cogito* improves on the ‘official story’ of deductive inference, see e.g. Ciprotti (2009), since it can explain a couple of problems identified by previous interpretations, as well as explain additional features that had not yet been identified (e.g., use of TE indicators, problems of validity and soundness, which this account can explain by transformation of the raw TE premise into a valid TE argument). The question why Descartes didn’t state the argument as a syllogism, modus ponens, himself may be answered in a rhetorical manner. If Descartes may have denied that it was a syllogism, he either wanted to hide his rhetoric, or he was hesitating to state it as deduction because it is firstly about an individual experience; or perhaps he was just overcorrect

should likewise cease entirely to exist . . . to speak accurately I am not more than a thing which thinks. (DESCARTES quote by RAHMAN, 1963, p. 488).

Of course, like Ibn Sinā’s *Flying Man* thought experiment, the *Cogito* thought experiment is only one argument for dualism in Descartes (and a comparatively weak one), and like in Ibn Sinā the immortality of the soul is another argument which has historically been considered stronger. However, it should be noted that strictly speaking the immortality of a transcendent soul cannot be inferred from the *Cogito*. In the objectivist phenomenal sense that may relate to traditional differences between, say, mind (or self) and soul, the *Cogito* thought experiment is similar to present day thought experiments in philosophy of mind, featuring on consciousness (studies).
for, in fact, ‘cogito ergo sum’ is an enthymeme or incomplete syllogism of which a part – in this case the major – is implicit. Handbooks of rhetoric may point out that statements are often excluded for strategical reasons when they are too obvious, and revealing them could damage the force of the argument, e.g.: “By making the reader work out the syllogism for himself, you impress the conclusion upon him, yet in a way gentler than if you spelled it out in so many words [...].” (HARRIS, 2010).

10. Applying terminology from present day thought experiment discussions, Descartes’ Cogito is a strong thought experiment because it generates new knowledge and insights (mind-body dualism, subjectivism), new empirical data as the I or subject (whose existence can be linguistically, sociologically, etc., that is empirically proven by its effects on external, objective reality as language, societal institutions, and perhaps also neurological effects) and because the general scope of the conclusion goes far beyond the particular situation of the thought experiment.7

4 Descartes’ Methodology of Doubt

The Cogito is the most transparent result of Descartes’ methodology of doubt, that, according to some, shows the influence of Al-Ghazali—instead of Ibn Sinâ’s Flying Man, a Persian thought experiment to which the Cogito is often compared in the history of philosophy.8 Descartes’ methodology of doubt is prominent in, e.g., the subtitle to Meditation I, ‘Of the things we can doubt’ (‘De iis quae in dubium revocari possunt’), and truth or certainty is defined as something ‘impossible to doubt’ (‘aliquid certi atque indubitati’). In our contemporary age, philosophers such as Daniel Dennett (1991, p. 3) return to Descartes’ Meditations, but instead of the Cogito, Descartes’ lesser-known Evil Demon thought experiment is revived, e.g., as the Brain in a Vat

7 The terms used here are derived from e.g. John Norton, Tamar Gendler, Thomas Kuhn, Roy Sorensen, James Brown a/o. See, e.g., Norton (1991, p. 129) ‘inductive step’: “Thought experiments are arguments which [...] invoke particulars irrelevant to the generality of the conclusion.” In terms of Thomas Kuhn (1962/1970, 1964) the Cogito is not only strong and paradigmatic, but a ‘revolutionary’ thought experiment.

8 See Hertogh (2015), in which it is demonstrated that the logical structures of Ibn Sinâ’s Flying Man and Descartes’ Cogito are too different from each other to suggest that Descartes may have known Ibn Sinâ’s thought experiment when he was conceiving of his Cogito thought experiment, in other words that it is very unlikely that Ibn Sinâ’s Flying Man may have inspired Descartes to his Cogito. Nevertheless, Descates may have known of Ibn Sinâ’s oeuvre and it may have inspired him in other places of his own oeuvre.

It is also in the *Philosophical Investigations* that we may find arguments not so much against Descartes’ methodological presupposition of the de facto possibility of introspection, but against Descartes’ near equivalence of introspection with thinking or cognition (res cogitans). On the basis of, e.g., #412-413, #551, #587, and p. 231-232, we could argue on behalf of Wittgenstein that introspection is rather like feeling than thinking, and that (Descartes’) thinking is an umbrella term for many mental processes, like feeling, imagining, memorizing, calculating, or merely ‘peculiar motions in the head and between the head and the throat’, etc., not all of which we could label exclusively cognitive. These are some of the reasons why we prefer a logical approach over a psychological one (there is too much ‘conceptual confusion’ in psychology) (WITTGENSTEIN, 1953, p. 232), aside from the fact that semantics can relate the thought experiment analyses to external reality.

**ACKNOWLEDGEMENTS:** This paper was developed by an elective seminar, Logic and Argumentation: Thought Experiments Capita Selecta, delivered at Chongqing University, Chongqing, China, in the first term of 2013-2014. Subsequent parts of the paper were presented as a lecture on *Semantics of Thought Experiments* for PhD Day of the Center for Logic and Philosophy of Science, Free University Brussels (VUB), on January 12, 2015, and as PPT to Private PhD Defense on September 10, 2015, VUB, Brussels. Thanks to the Library Learning Center of University of Amsterdam (UvA), Amsterdam, for assistance in preparation.


**RESUMO:** O *Cogito* de René Descartes é um exemplo de experimento mental paradigmático, precursor tanto do subjetivismo quanto da nova ciência, na Idade Moderna europeia, o qual parece ter escapado à atenção dos filósofos que estudaram o experimento mental. Na análise profunda, o *Cogito* aparece como uma instânciação universal (ou *modus ponens*, com implicita principal “tudo o

---

9 Wittgenstein (1953, #413): Here [see #412] we have a case of introspection, not unlike that from which William James got the idea that the ‘self’ consisted mainly of ‘peculiar motions in the head and between the head and throat’. And James’ introspection shewed, not the meaning of the word ‘self’ (so far as it means something like ‘person’, ‘human being’, ‘he himself’, ‘I myself’), nor any analysis of such a thing, but the state of a philosopher’s attention when he says the word ‘self’ to himself and tries to analyse its meaning. (And a good deal could be learned from this.)
que tem a propriedade de pensamento existe”). O Cogito tem fortes efeitos retóricos por si mesmo, generalizando narrativamente desde o eu para toda a espécie humana, e seu sucesso histórico e filosófico pode ser explicado por sua estrutura entinemática concisa, que soa através de muitos sentidos possíveis. Consideramos que é um exemplo proeminente de um experimento mental, na medida em que afirma o poder de pensar como seus próprios conteúdos. A partir da metodologia da dúvida de Descartes, podemos concluir que (por exemplo), em uma interpretação wittgensteiniana, o Cogito é um experimento mental mais lógico que psicológico.


REFERENCES


HERTOGH, C. P. Turn to Butterfly dream. Presentation delivered on May 15, 2013, to Prof. Weixue Li’s course on Dao De Jing. Chongqing, Chongqing University, 2013. [PPT]


