

DESCARTES AND HUME ON I-THOUGHTS

DESCARTES Y HUME ACERCA DE LOS PENSAMIENTOS SOBRE EL YO

Luca Forgione¹
Universidad de Basilicata (Italia)

Recibido: 29/09/2017

Aceptado: 4/01/2018

Abstract: Self-consciousness can be understood as the ability to think I-thoughts which can be described as thoughts about oneself ‘as oneself’. Self-consciousness possesses two specific correlated features: the first regards the fact that it is grounded on a first-person perspective, whereas the second concerns the fact that it should be considered a consciousness of the self as subject rather than a consciousness of the self as object. The aim of this paper is to analyse a few considerations about Descartes and Hume’s approaches to self-consciousness, as both philosophers introduce a first-personal method of accessing the subjective dimension through an introspective account. Descartes’s view on self-consciousness seems incapable of conceiving and recognizing herself as herself, while Hume’s seems to lack those features assigned to the consciousness of self-as-subject.

Keywords: Self-consciousness, I-thoughts, first-person perspective, Descartes, Hume.

Resumen: La autoconciencia puede ser entendida como la habilidad para llevar a cabo reflexiones sobre uno mismo, las cuales pueden ser descritas como los pensamientos sobre tu propia persona como “individuo”. La autoconciencia comprende dos características relacionadas entre sí: la primera tiene que ver con el hecho de que está basada en una perspectiva en primera persona, mientras que la segunda tiene que ver con el hecho de que debe ser considerada como la conciencia del yo como sujeto en lugar de como objeto. El objetivo de este trabajo es analizar algunas cuestiones acerca del punto de vista de Descartes y de Hume sobre la autoconciencia, ya que ambos filósofos introducen un método de acceso en primera persona a la dimensión subjetiva a través de un

1. Associate professor of philosophy of mind and language at the University of Basilicata (Italy).
Email: luca.forgione@gmail.com

enfoque introspectivo. La visión de Descartes sobre la autoconciencia parece incapaz de concebirla y reconocerla como tal, mientras que la de Hume parece carecer de las características asignadas a la conciencia como sujeto.

Palabras clave: autoconciencia, filosofía, perspectiva en primera persona, Descartes, Hume

1. First-person perspective and the consciousness of the self as subject

The first-person perspective point can be explained by Baker (2000, 2013)'s approach. Two types of first-person perspectives can be distinguished: a *rudimentary first-person perspective*, manifested by many mammals and human infants, and a *robust first-person perspective*, manifested by language users who master first-personal language. The latter is the conceptual capacity not only to recognize oneself as distinct from things other than oneself, but also to conceive oneself as oneself. A robust first-person perspective is what exactly marks the difference between a creature with a rudimentary first-person perspective who can only be conscious of the environment, and a fully self-conscious subject. As a matter of fact, a mature human subject with a robust first-person perspective can attribute to herself a first-person reference on the basis of a self-concept, i.e., not only can she refer to herself in the first-person, but she can also attribute first-person reference to herself.

A crucial distinction is made between making first-person reference (as when Pasquale says, "I am tall") and attributing first-person reference (as when Mario says, "Pasquale wishes that he himself were tall"). With the latter case, Mario attributes to Pasquale a wish he would express by a first-person reference. The attribution of a first-person reference occurs in indirect discourse, in a "that"-clause following a psychological verb. The point is that a subject does not attribute first-person reference only to others but also to his own self, as when Pasquale says, "I wish that I were tall". A subject thinking "I am tall" can distinguish herself from others; a subject thinking "I wish that I were tall" can conceptualize that distinction: she can think of herself as herself. This ability to attribute first-person reference to oneself is the manifestation of strong first-person phenomena.

Following Baker (2013) and Matthews (1992), "I*" pronouns² are used reflexively to pick out the subject from her own point of view: gi-

2. Castañeda (1966; 1967) employs an asterisk, or star, next to a pronoun ("he*") to attribute first-person reference to someone else, as in "Pasquale believes that he* is tall. This sentence is not true unless Pasquale expresses his belief in the first person: "I am tall." Matthews (1992) introduces the "I*" for sentences with first-person subjects in order to analyse the

ven the close relation between the linguistic and the mental dimensions, I*-sentences are sentences containing “I*”, whereas I*-thoughts are thoughts expressible by I*-sentences. By an I*-thought a subject conceives herself as herself*, and needs no third person referential device, such as a name, description, or demonstrative to identify herself. As we will see in the paragraphs 2 and 3, certain semantic and epistemic features of the term/concept *I*³ can be identified in this subject’s capacity of self-identification: essential indexicality and immunity to error through misidentification. The former is relative to the meaning of the term/concept *I*, any expression of self-awareness being based on indexical terms such as “I” or “me”; the latter, on the other hand, refers to the fact that some singular judgments involving the self-ascription of mental (and physical, as will be seen) properties are immune to error through misidentification relative to the first-person pronoun (IEM). The subject formulating such judgments in given epistemic contexts cannot be mistaken as to whether it is he himself who is attributing a particular mental property to his own self.

At the same time the basic self-consciousness at issue here is also to be regarded as the consciousness of self-as-subject, or *subject self-awareness*, rather than the consciousness of the self as object. Following Kriegel (2003, 2007), it is possible to make a distinction between the consciousness of oneself *qua* object and the consciousness of oneself *qua* subject. For instance, Mario can be conscious of Naples: Mario is the subject of the thought, and Naples its object. Mario, however, can also be conscious of himself*: in this case, Mario is both subject and object of the thought. Even though there is one single entity, and the subject and object of the thought are the same thing, it is possible to draw a conceptual distinction between Mario’s ability as object of thought and Mario’s ability as subject of thought: namely, the concepts of self-as-subject and self-as-object. The self as subject may be thus interpreted as the thing that does the thinking, whereas the consciousness of oneself as subject is the consciousness of oneself as the thing doing the thinking.

It seems clear that the first-person perspective and the consciousness of self as subject are two interdependent features, the one being the condition of the other and *vice versa*. The subject’s manifestation of strong first-person phenomena – one’s ability to attribute first-person reference to herself – is based on the subject’s possibility of being the consciousness of herself as the thing doing the thinking, and the subject’s consciousness

phenomena expressed by “I think that I* am F”.

3. I have developed the issue in XXXXX (2017).

of self-as-subject cannot be gained unless the subject exhibits a manifestation of strong first-person phenomena.

As has been just said, such two features defining the notion of basic self-consciousness are grounded on a few epistemic and semantic peculiarities in the ability to use the term or concept *I* in *de se* or *I-thoughts*: the essential indexicality and the immunity to error through misidentification. These will be discussed in paragraphs 2 and 3.

2. Indexicality

As has already been said, an I*-thought allows an individual to refer to herself as herself* without a need for third-person referential devices, such as names, descriptions, or demonstratives, to identify herself: the term/concept 'I' employed in a self-conscious or I*-thought is essentially indexical⁴; as such, it necessarily involves information indexed to the context and – more specifically – to the thinker who has produced the thought. In detail,

(a) 'I' is a singular term/concept, that is, a term with a single individual as its reference;

(b) this term is governed by the token-reflexive rule, whereby every token of 'I' refers to the subject who has produced or used it, either mentally or linguistically;

(c) with the information available in context, and once the circumstances of evaluation are established, *prima facie* this rule is sufficient to determine its reference⁵.

More importantly, the indexical information about oneself based on the use of the term/concept 'I' cannot be reduced to non-indexical information; for this reason, indexicality is *essential*. Two well-known examples by Perry (1977) describe the matter at issue.

The first example regards a fictional character named Rudolf Lingens: “An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost. [...] He still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say, *This* place is aisle five, floor six, of Main Library, Stanford. *I* am Rudolf Lingens”. Amnesiac Rudolf Lingens can gather all sorts of information about himself by reading the

4. Cf. Kaplan (1989) and Perry's (1997, 594) classic approaches.

5. This approach to indexicality has been acknowledged by major scholars. Nonetheless, this position is not entirely uniform, cf. de Gaynesford (2006) for an analysis of the debate.

books in the Library, and yet no such information can provide him with the missing conceptual tool he needs to link the information with himself. In other words, there is no logical connection between third-person descriptive information, no matter how detailed, and a first-person grasp of oneself through the use of 'I'.

The second example regards indexical judgements (beliefs and desires), which are crucial to explain and predict the motivating action: "I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch" (Perry, 1979).

In the example, Perry's thoughts (a) "the shopper with the torn sack is making a mess" and (b) "I am making a mess" refer to the same individual, so two intrinsically different kinds of self-reference are at play here. In the former, self-reference is external and available in the third person only: Mario can refer to an object by using a name, a definite description or a demonstrative, and the object he is referring to might be himself; there is no difference between this kind of self-reference and the reference made to an object that is different from oneself. An external self-reference can occur without the subject's realizing he is referring to himself, as in the first thought formulated by Perry, where he does not realize he is the very shopper with the torn sack who is making a mess.

Instead, the internal self-reference expressed by the second thought (b) "I am making a mess" produces an authentic I-thought that is only accessible from a first-person perspective because it is based on the use of 'I'. As soon as he realizes that he is the shopper at issue, Perry produces a new thought, which he terms *locating belief*, based on the use of the essential mental indexical 'I'. This entails that the ascription of an authentic I-thought to one's own self cannot be achieved without the concept 'I' as there is no way to think an I-thought other than through indexical reference; hence, I-thoughts are irreducible to the other kind of non-indexical thoughts.

Accordingly, by articulating an I-thought in a propositional way, an I-thought will contain a content whose subjective essential indexical reference is expressible in the natural language by the personal pronoun 'I'. This thought can be reported in either direct or indirect form: while the former is in *oratio recta* and reports the above-mentioned example as in (1) "I am making a mess" (thought by Perry), in the indirect form the thought can be expressed *in oratio obliqua*, and the report will be in the third person: (2) "Perry thinks that he himself is making a mess". In turn, this sentence can be interpreted as the report of yet another thought still: (3) "Perry thinks

that Perry is making a mess”. Obviously, it is possible to employ a definite description, “the ϕ ” that picks out Perry uniquely, as in this example: (4) “Perry thinks that the author of *The Essential Indexical* is making a mess”.

The thought expressed in (1) is neither equivalent to (3) nor (4). With cases (3) and (4), Perry might be amnesiac and neither remember his name nor his being the author of *The Essential Indexical*. It is only in (1) that an authentic I-thought is present: the subject who thinks the thought “I am making a mess”, provided that she knows the rule associated with ‘I’, cannot use it without realizing that she is referring to nobody but herself. Although (1) and (3) or (4) are not equivalent, *prima facie* (2) seems to be a report of both. To capture the difference, Castañeda employs two different uses of the third person pronouns in *oratio obliqua* sentences. In the first case, to make (2) equivalent to (1), the pronoun is to be used in an *indirect reflexive* modality (Anscombe, 1975) or as a *quasi-indicator* – the above-mentioned artificial pronoun (‘she*’, ‘he*’, ‘it*’) introduced by Castañeda (1966, 1967, 1968) to attribute a first-person essential indexical use from a third-person angle: (2.1) “Perry thinks that he* himself is making a mess”. The *quasi-indicator* “he*” in the example is used as an anaphora, and its reference is not determined directly but only through the propositional attitude subject. In the other case, assuming that Perry is amnesiac, to make (2) equivalent to (3) “he” will not be employed as a quasi-indicator but as a simple indexical; thus, (2) is the report of Perry’s belief that someone else in the context (named Perry too) is making a mess: Perry has not realized that it is he* himself who is doing that.

What matters here it’s the fact that token-reflexive expressions such as first-person pronouns and quasi-indicators are essential indexicals: they can be neither eliminated nor replaced by a name, description, or demonstrative without losing the content expressed by the sentences/ thoughts that contain them: to refer to (to think of) oneself *qua* oneself, the subject has to use the essential indexical ‘I’. Therefore I-thoughts make up an irreducible class of mental phenomena.

3. Immunity to error through misidentification

In a well-known passage, Wittgenstein (1958, 66-7) introduces his philosophico-linguistic analysis of the grammatical rule of the term I, where he identifies two types of uses, i.e., ‘I’ used as object (“I have grown six inches”) and ‘I’ used as subject (“I have toothache”): “One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these

cases the possibility of an error... On the other hand there is no question of recognizing a person when I say I have toothache. To ask ‘are you sure it’s you who have pains?’ would be nonsensical”.

This passage should be taken as part of the philosophical framework articulated by Wittgenstein since the 1930s according to some theses (some of which will be examined in Chapter 4) to be regarded as the background for the analyses of the two uses of ‘I’. While the *I* used as object performs a referential function relative to one’s body and physical features in general, the *I* used as subject apparently regards mental states and processes involving no subject identification.

Particular judgments displaying first-person reference (e.g., “I have pain”) display what Shoemaker (1968, 565) defines *self-reference without identification*: “My use of the word ‘I’ as the subject of my statement is not due to my having identified as myself something of which I know, or believe, or wish to say, that the predicate of my statement applies to it”. The self-ascription of the thoughts on which the self-consciousness is based regards the consciousness of oneself *qua* subject – namely, as the subject of every thought or mental state – rather than as the object based on the previous identification. Due to the absence of any identification component, particular singular judgments involving the self-ascription of mental (and physical, as will be seen) properties are *immune to error through misidentification* relative to the first-person pronoun (IEM). For example, if a subject feels pain and judges ‘I am in pain’, the subject formulating such judgments cannot be mistaken as to the person who happens to be in pain.

To make this point clear, we might consider an example of identification-dependent thought, e.g., Babis’s thought that his neighbour is a nice person. Following Evans (1982) and Kriegel (2007), the structure of this thought consists of an identification component and a predication component, which can be explained by Babis’s first-person perspective as follows: “my neighbour [identification component] is a person who smiles at me every day and the person who smiles at me every day is a nice person [predication component]”. Here, two types of errors are possible. Babis can be mistaken as to the predicational component, i.e., that his neighbour is a nice person: for example, later on Babis finds out that his neighbour’s tendency to smile is nothing more than a cynical strategy to have him consent to cut the trees in the garden. Babis can also be mistaken as far as the identificational component is concerned – i.e., with respect to the person who is his neighbour – and, for example, get confused and mistake the mailman for his neighbour.

On the other hand, there is a class of self-ascriptions involving no identification-dependent thoughts; as such, it is not subject to error through misidentification. Shoemaker (1968, 565) examines the kinds of

psychological predicates involved in such self-ascriptions: “There is an important and central class of psychological predicates, let us call them P^* predicates, each of which can be known to be instantiated in such a way that knowing it to be instantiated in that way is equivalent to knowing it to be instantiated in oneself”. For instance, the judgment “I have pain” is IEM because the way in which the predicate is expressed (“there is pain”), that is, based on our own subjective experience, will suffice to realize that it is ascribed to ourselves (“I have pain”). It is in this particular sense that “there is pain” is tantamount to “I have pain”.

Evans, in turn, goes beyond the terms of the matter as suggested by Wittgenstein and, to some extent, by Shoemaker. In particular self-ascriptions, the self-reference is direct and unmediated: as Evans notes, this is *identification-free self-reference*. More to the point, and moving from the self-ascription of properties that are not only mental but also physical, the author discloses his approach: judgments are IEM when they result from the connection between the information acquired in the first person and its justification, as opposed to *identification-dependent* judgments involved in the ordinary perception of external objects. The IEM feature does not depend on the kind of predicate involved in the self-ascription but on the epistemic and justification ground on which the subject produces such judgments in a context where – from Strawson’s lesson onward – the subject is conceived as a spatio-temporally located object.

In particular, Evans (1982, 220) contends that a judgment such as “I am F” is identification-free unless it corresponds to the inferential conclusion drawn from the two premises, i.e., “a is F” (predication component) and “I am a” (identification component). Such a judgment is based on the unmediated self-ascription of properties through the introspective consciousness (as is the case with mental properties) or proprioception (as with physical properties). For example, according to our general capacity to perceive bodies and to our sense of proprioception, of balance, of heat and cold, and of pressure, the kind of information generated by each of these modes of perception seems to give rise to immune to error through misidentification judgments: “None of the following utterances appears to make sense when the first component expresses knowledge gained in the appropriate way: ‘Someone’s legs are crossed, but is it my legs that are crossed?’”.

Therefore, the self-consciousness capacity depends on the possibility to produce I-thoughts, which employ indexical self-reference immune to error through misidentification relative to the term/concept *I*. As said before, paragraph 4 will examine some considerations on Descartes and Hume’s views on self-consciousness as both of them introduce a first-personal method of accessing the subjective dimension through an introspective account. While Descartes’s view of self-consciousness seems unable

to conceive and recognize oneself as oneself*, Hume's seems to lack the features assigned to the consciousness of the self as subject.

4. Descartes and Hume on I-thoughts

In *Meditations on First Philosophy*, with his famous *cogito* argument Descartes introduces an introspective account of self-awareness. Having employed the methodological and rational doubt to be sceptical about all kinds of knowledge and opinions commonly held to be true, Descartes does not build his argument in its standard formulation (*cogito ergo sum*), but through certain logically equivalent steps aimed at demonstrating that, even when in doubt, the subject cannot doubt about the doubting itself, as she cannot question the existence of the thinking that manifests itself in the mental sphere.

The verb *cogitare* and the mentalist verbs employed are in the first person because the existence of the thinking is only undoubted from a first-person perspective. Since the subject is conscious of her own thoughts, she can conquer this certainty inside her mental sphere: whenever a mental activity is given, and whatever its nature (hoping, doubting, willing, etc.), the subject captures it as well as its existence through intuition to specify the introspective account of self-awareness by means of an acquaintance method⁶.

In metaphysical terms, thought is the essential nature of the mind: "I find here that thought is an attribute that belongs to me; it alone cannot

6. Going through the steps that outline the Cartesian distinction between deduction, i.e., the inference of something that necessarily follows from other propositions that are known with certainty, and intuition, i.e., the faculty by which it is possible to capture the initial certainty of a proposition in an immediate and self-evident way to make deduction possible, Markie (1992, 145) asserts that in the *cogito* argument Descartes's use of intuition is based on two different types of intuition: "The set of intuited propositions includes both self-evident propositions not inferred from any others and propositions immediately inferred from self-evident premises". Propositions about mental states are self-evident and not inferred from others – i.e., they are intuitive in a narrow sense – while propositions about one's existence are immediately inferred from propositions about mental states, i.e., they are intuitive in a broad sense. In support of this argument the author shows several Cartesian passages (e.g., his replies to Mersenne's objections) where he refers to *lumen naturale*, *ratio naturalis*, and clear distinct perceptions as criteria to explain his certainty about the thinking and its existence. There are, however, several problems related to *The Self-Evident Intuition/Immediate Inference Interpretation*: Markie himself reformulates his argument to reconcile the contrast between the passages where Descartes explicitly refers to the two senses of intuition as the grounds of the *cogito* argument and those claiming the need for prior knowledge of the general proposition for which it is impossible for what thinks not to exist (for example, cf. the tenth article of the first part of the *Principles of philosophy*). A syllogistic structure of the *cogito* argument is thus maintained (cf. Williams, 1978). Other commentators bring into question

be separated from me. I am, I exist, that is certain. But how often? Just when I think” (Descartes 1641, 51-2). The Cartesian approach is based on the shift from an epistemic to a metaphysical argument; on the one hand, the former concerns how the subject represents herself as a thinking entity in the mental sphere. In the Twentieth century this has sometimes been discussed – in particular to be dismissed – as a form of privileged access (cf. *supra*): each subject accesses her mind in an infallible way compared to how she accesses external objects, including her own body. The second argument, on the other hand, identifies two distinct substances: *res cogitans* and *res extensa*⁷. In this manner Descartes (1641, 20-1) assigns a special role to self-awareness to achieve his metaphysical results:

But what therefore am I? A thinking thing. What is that? I mean a thing that doubts, that understands, that affirms, that denies, that wishes to do this and does not wish to do that, and also that imagines and perceives by the senses. Well, indeed, there is quite a lot there, if all these things really do belong to me. But why should they not belong to me? Is it not me who currently doubts virtually everything, who nonetheless understands something, who affirms this alone to be true, and denies the rest, who wishes to know more, and wishes not to be deceived, who imagines many things, even against his will, and is aware of many things that appear to come via the senses? Is there any of these things that is not equally true as the fact that I exist—even if I am always asleep, and even if my creator is deceiving me to the best of his ability? Is there any of them that can be distinguished from my thinking? Is there any that can be said to be separate from me? For that it is I that am doubting, understanding, wishing, is so obvious that nothing further is needed in order to explain it more clearly.

This passage reveals that the various mental activities (doubting, understanding, etc.) belong to a thinking substance that should be understood as an I (“it is I that am doubting, understanding, wishing”): there is no mental activity without an *ego*, i.e., without a subject performing such

the very possibility of inferring the existence of the subject from the consciousness of thought, cf. Kenny (1968, 169), Wilson (1978, 55), and, obviously, Hintikka (1962). For an effective critique to Hintikka’s non-inferential reading of the *cogito* argument, cf. Bonomi (1991, 19): his reading is similar to Markie’s (1986, 1992) in some respects, especially when he focuses on the particularly plain inferences highlighted by Descartes, also regarding the *cogito*, that can be grasped intuitively to affirm the compatibility between the inferential nature of the *cogito* and its intuitive character.

7. Cf. Kenny (1968), Hooker (1978, 171-4) and his *Descartes’s Argument from Doubt*. Having gained the certainty of the *cogito*’s existence, the criterion employed by Descartes to assume an ontological dualism is, so to speak, external: once its existence is demonstrated, in the sixth meditation Descartes refers to God and relies on the fact that he will not deceive about the clear and distinct perception of the self as a thinking and unextended substance, establishing the intrinsic property of the mind which is ontologically different from the extended body; the immaterial essence of the *ego* is thus identified.

an activity⁸. The metaphysical connection between the I and the thoughts is nothing but an example of the metaphysical relationship between a substance and its attributes. An argument challenged by Hume, as will be seen further on. As a matter of fact, in his reply to one of Hobbes's objections, Descartes (1641, 110) points out that "it is certain that thought cannot exist without a thinking thing, nor can any act or any accident at all exist without a substance in which it inheres".

If the link between a substance and its attributes is metaphysically obvious, then the relation between a thinking substance understood as *ego* and the thoughts from which the Cartesian approach to self-consciousness can be characterized is far more complex. When the subject turns her gaze inwards, that is, to the mental sphere, what she grasps through the conscious dimension are thoughts, i.e., the attributes of the substance she is thinking of: generally speaking, if a substance cannot be captured in itself, but always and only through the attributes that are inherent in it, when a subject turns her gaze toward the mental sphere she does not capture a substance in itself, but only the thoughts that are inherent in the thinking substance (cf. Markie 1992, 171; Wilson 1978, 66).

If, on the one hand, this plainly means that to be self-conscious one needs appeal to the metaphysical sphere and to the *res cogitans* itself, on the other – and less plainly – the only reference to the metaphysical thinking substance will not be as sufficient a condition: it will not suffice to justify Descartes's certainly innovative use of the term 'I' to refer, as a thinking substance, to himself *as himself*^{*} without any conceptual mediation ("it is I that am doubting, understanding, wishing, is so obvious that nothing further is needed in order to explain it more clearly").

Beyond the objection made by Lichtenberg in the Eighteenth century, that what Descartes can actually state is "there is a thought", rather than "I'm thinking" (at the time, an objection resumed by logical empiricists such as Mach and criticized by Williams, 1978), it is exactly in the

8. And there is no thought without an awareness of thinking: the consciousness is an intrinsic feature of thought. The argument is present in Locke but is neither implied in Leibniz nor in Kant, who employs Leibniz's notion of *petits perceptions* through Wolff and Meier's *dunkle Vorstellungen* to assume the possibility of unconscious mental representations. Although a few commentators have tried to soften the rigid Cartesian identification of thought and consciousness by referring to other passages of his works (cf. Gombay 2007, 122-3), it is in the context of the *Meditations on First Philosophy*, if nowhere else, that Descartes seems straightforward on this point: when he replies, for example, to Arnauld's objections, "there can be no thought in us of which, at the moment at which it is in us, we are not conscious" (1641, 158). This point had already been stated earlier in his replies to the second objections: "I use the term thought to cover everything that is in us in such a way that we are immediately conscious of it" (1641, 102).

terms of self-awareness that the link between the thinking substance understood as an I and the thoughts seems to be problematic.

A subject captures the thoughts in the mental sphere, and this may be properly reported by the statement “the thought is ongoing”⁹. As a consequence, the subject can only come to the conclusion that “the thing with *this* thinks”, where “this” refers to the thought in course rather than to “I think”. Nonetheless, a full self-conscious capacity can be expressed only through first-person assertions such as those used by Descartes: the question is how Descartes, beside a few sporadic references to the concept ‘I’, justifies the fact that thoughts captured in the conscious sphere are not the attributes of some thinking substance but *his* very attributes, which can be related to a particular thinking substance which turns out to be *himself*.

In this respect the Cartesian approach does not offer an explicit characterization of the indexical thought in order to constitute *I*-thoughts* or *de se* assertions¹⁰, nor does it put forward an exhaustive explanation of the subject’s ability to identify himself *as himself* based on the mere presence of the thought summarized by the statement “the thing with this (thought) thinks”. Markie (1992, 164) offers the following example to clarify the Cartesian difficulties:

Suppose that he [Descartes] considers one of his ideas, does not yet know whether he has produced it or whether God has produced it in him, and decides that whatever has produced it is perfect. Suppose too that he is the source of the idea. Clearly, the thought that Descartes would express by “The thing that produced this is perfect”, where “this” refers to the idea, is not the same as the one he would express by “I am perfect”. He believes the former but he may not believe the latter. The difference between the thoughts is that, although Descartes thinks of himself in each thought – he is the referent of both “the thing that produced this” and of “I” – in the first thought he only thinks of himself and in the second he thinks of himself *as himself*. This difference between the two thoughts is lost, if we analyze the thought Descartes would express by “I am perfect” as the one he would express by “The thing that has this is perfect”, where “this” again refers to the idea.

In the first book of his *Treatise of Human Nature*, Hume (1739-40, 364) dismisses the idea that the subject can perceive a persistent self

9. This suggests yet another argument still for a syllogistic interpretation of the *cogito*: from the general principle for which any observed quality belongs to a substance, and given a thought caught in the conscious sphere, the subject infers that a thinking substance does exist. Nonetheless, Markie (1992) develops this argument in the terms of his *The Self-Evident Intuition/Immediate Inference Interpretation* through a few epistemological distinctions. Besides, on several occasions Descartes himself argues that the *cogito* is not based on a syllogism: obviously, as Markie and Bonomi’s interpretations suggest, the non-syllogistic nature of the *cogito* does not imply its non-inferentiality.

10. The debate on Descartes and self-reference is obviously vast and offers different positions, e.g., Anscombe (1975), Van Cleve (1983), Zemach (1985), Markie (1986, 1992).

through the introspective consciousness to eventually suggest an elusive epistemic thesis:

For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch myself at any time without a perception, and never can observe any thing but the perception.

Hume's approach fits into the framework of a science of human nature which is grounded on both the experimental method and the Newtonian principles to explain the genesis of mental events and the association of ideas according to the principles of resemblance, contiguity, and causation. Perceptions are divided into impressions and ideas, simple and complex, deriving from sensation and reflection. While impressions of sensations cause ideas – which are to be distinguished in their being weaker images of impressions – ideas, in turn, cause impressions of reflexion: passions, desires, and emotions.

Hume tackles a form of introspective self-consciousness based on what Locke terms “inner sense” (cf. *supra*), its operations being, *mutatis mutandis*, similar to those involved in the perception of external objects. He challenges the assumption of the subject's encountering the idea or impression of herself in the flow of psychological perceptions in much the same way she encounters the idea or impression of red in the perception of the external world (cf. Stroud, 1977, 118). Such a result is chiefly reached to criticise the philosophical notion of substance: perceptions should be conceived as separable, distinguishable, and different – to Hume, they can perfectly exist without something like a substance to support them. Failing the notion of substance and the associated distinction between substance and accidents, the metaphysical conception of the self as substance with thoughts as its accidents dissolves, to be replaced by the well-known theory of the mind as a bundle of perceptions.

With Chisholm (1976, 38), the Humean thesis is articulated as follows: 1. I cannot be directly aware of any object unless that object is an impression; 2. but I am not an impression; 3. therefore I cannot be directly aware of myself. The first premise stems from his criticism of the notion of substance and the distinction between substance and accident: to Hume, one can only deal with mental particulars, impressions and sensations, as no idea of substance is distinct from that of a collection of particular qualities. If an impression is required for an idea to be produced, then, according to the second premise, no impression of an ego can form its idea: “self or

person is not any one impression” (Hume, 1739-40, 263)¹¹. If there were any such thing as an impression attributable to a self, then at least this should be constant: assuming that the self exists, it must exist in a constant, invariable way. To Hume, however, in the flow of perception no impression stays unaltered, least of all when it relates to a self. It follows that if the subject is only conscious of impressions, and if no impression is related to the self, then the subject cannot be conscious of herself in the flow of perception.

At the same time, we may reach a metaphysical thesis of *exclusion*: the thinking subject is no substantial object in the world. This thesis is obtained as a conclusion of the previous *elusive* epistemic thesis: given that, in introspection, the thinking subject cannot manifest itself to itself nor can be known as an object (epistemic argument), then the subject is no substantial object (metaphysical thesis).

From a metaphysical perspective, what is denied here is the existence of a substantial ego, although this will not affect the very possibility of self-consciousness, as pointed out by Frank (2004, 71): “Hume failed to realize that self-consciousness and ego-consciousness are not the same, so that the critique of Descartes’ substantialisation of the ego in no way shows the possibility of a knowledge of consciousness, and thus of self-consciousness, to be incoherent”. Firstly, it is precisely through self-consciousness that the eliminative conclusion about the substantial ego can be attained: after all, it is exactly when he most intimately enters what he refers to as “myself” that Hume comes to the conclusion of his merely stumbling on a few particular perceptions¹². Secondly, having rejected a substantialist perspective of the self upon which the psychological flow of thoughts ultimately depends, Hume (1739-40, 263-4) puts forward the famous argument of the bundle of perceptions:

I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement. [...] The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations. There is

11. Here it is already possible to bring up the vicious circularity objection proposed by Noonan (1999, 193), who questions how Hume can take for granted that the self is no impression whatsoever, since he does not know what it really is: “If Hume has no idea of a self he presumably has no conception of what it would be like to observe one. In that case, however, how does he know that he is not doing so? Maybe he is, but just fails to recognize the fact”.

12. Chisholm (1976, 40) also finds a problem in this particular argument by Hume: “As Professor Price once observed, it looks very much as though the self that Hume professed to be unable to find is the one that he finds to be stumbling – to be stumbling on to different perceptions. How can he say that he doesn’t find himself – if he is correct in saying that he finds himself to be stumbling and, more fully, that he finds himself to be stumbling on certain and not to be stumbling on certain other things?”.

properly no simplicity in it at one time, nor identity in different.

The thesis is entirely built on the reification of perception, understood as an entity independent of any substance from both a logical and ontological point of view, as remarked by Noonan (1999, 197):

the relation between the self and its perceptions is analogous to that between the sea and its waves. The waves are modifications of the sea and perceptions are modifications of the self. But Hume, in claiming that perceptions are logically ontologically independent, denies this and thus denies the only possible basis for regarding the self, *qua* perceiver, as ontologically prior to its perceptions. That he should claim that the self is in reality nothing but a bundle of its perceptions [...] is thus entirely intelligible. Once perceptions are reified as substances no other conception of the self makes any sense at all.

Beyond the issues concerned with circularity, self-awareness is understood as the exercise of a mind regarded as a bundle of perceptions producing the illusion of a self, the fictitious product of the flow of perception¹³. Hume's approach to self-awareness is based on an intrinsically observational, introspective account and, as such, is dismissed by those who regard basic self-consciousness as the consciousness of oneself as subject. This introspective account cannot account for the most fundamental type of self-consciousness – i.e., subject self-consciousness – but only object self-awareness at best. If an introspective account of self-awareness identifies the 'I' with the subject *qua* thing known through the properties observed within the introspective awareness, rather than with the subject *qua* knower, or introspector, then the subject can only be seen as the *object* of awareness, not as the thing doing the thinking.

It is worth noting that both Descartes and Hume's approaches are based on an introspective account of self-awareness; however, while Descartes employs an acquaintance account of introspection, Hume relies on an observational explanation. Despite the important difference in the way

13. This does not remove another difficulty that Hume acknowledges not being able to surmount: how to justify the genesis of this illusion according to given psychological rules. In the appendix to the third book of the *Treatise* dated 1740, Hume claims his inability to reconcile the principles that all our distinct perceptions are distinct existences and that the mind never perceives any real connexion among distinct existences, failing to provide an adequate explanation for the genesis of the idea of a self understood as a simple and identical reference; cf. Stroud (1977), Noonan (1999). At the same time, this problematic picture contrasts with the analysis of passions and morality in the second and third books, where the role of the ego is established without any uncertainty, given its importance for the explanation of the moral action, among others.

the subject introspectively reflects on her own thoughts in the two accounts, according to the core conception of introspectivism, basic self-consciousness is the consciousness of oneself as the thing with the thought being introspected – namely, as the thing with *this* thought. In other words, thanks to the introspected thought the introspecting subject recognizes herself as the owner of the introspected thought or as the thing with the thought she introspects. In the following passage, Shoemaker shows how an introspective account of self-awareness presents important difficulties, for the reason that it cannot explain how the subject cannot misidentify himself, i.e., how ‘I’ judgments are immune to error through misidentification (although his objection is mainly addressed to the inner sense account, his argument could also be used against Descartes’s approach, cf. *supra*):

[I]f the supposition that the perception [of my properties] is by “inner sense” is supposed to preclude the possibility of misidentification, presumably this must be because it guarantees that the perceived self would have a property, namely, the property of being an object of my inner sense, which no self other than myself could (logically) have and by which I could infallibly identify it as myself. But of course, in order to identify a self as myself by its possession of this property, I would have to know that I observe it by inner sense, and this self-knowledge, being the ground of my identification of the self as myself, could not itself be grounded on that identification. (Shoemaker 1968, 562–563)

In conclusion, the consciousness of oneself as the thing with “this thought” cannot be seen as basic self-consciousness: it depends on a subject’s prior self-awareness, which the introspective account does not manage to explain – what Shoemaker defines “the ground of my identification of the self as myself”.

References

- Almog J., Perry J., Wettstein H. (eds) (1989), *Themes from Kaplan*, Oxford U.P., Oxford.
- Anscombe, G. E. M. (1975), *The First Person*, in Cassam (1994), pp. 140-159.
- Baker, L. R. (2000), *Persons and bodies: A constitution view*, Cambridge U.P., Cambridge MA.
- ID. (2013), *Naturalism and the First-Person Perspective*, Oxford U.P., New York.
- Bonomi A. (1991), “Sul cogito cartesiano: natura inferenziale e cri-

teri di giustificabilità”, in G. Usberti (a c. di), *Problemi fondazionali nella teoria del significato*. Firenze: Olschki Editore, pp.7-33.

Cassam, Q. (ed.) (1994), *Self-Knowledge*, Oxford U.P., Oxford.

Castañeda, H.-N. (1966), ‘He’: A Study in the Logic of Self-Consciousness, in Id. (1999), pp.35-60.

Id. (1967), *Indicators and Quasi-Indicators*, “American Philosophical Quarterly” 4, 85–100.

Id. (1968), *On the Phenomeno-Logic of the I*, in Id. (1999), pp.89-95

Id. (1999), *The Phenomeno-Logic of the I. Essay on Self-consciousness*, Indiana U.P., Bloomington.

Chisholm, R. (1976), *Person and Object: A Metaphysical Study*, Open Court, Chicago and La Salle.

Id. (1981), *The First Person*, University of Minnesota Press, Minneapolis.

Descartes, R. (1641), *Meditations on First Philosophy. With Selections from the Objections and Replies*. Translated with an Introduction and Notes by Michael Moriarty. Oxford: Oxford University Press, 2008.

De Gaynesford, M. (2006), *I: the meaning of the first person term*, Oxford U.P., Oxford.

Evans, G. (1982), *The Varieties of Reference*, Oxford U.P., Oxford.

Frank, M. (2004), “Fragments of a History of the Theory of Self-Consciousness from Kant to Kierkegaard”. In *Critical Horizons* 5 (1): 53-136.

Forgione, L. (2017), *Self-consciousness and first-person perspective*. In: M. Curado, S. Gouveia (eds) *Philosophy of Mind: Contemporary Perspectives*, Cambridge: Cambridge Scholars Publishing.

Geach, P.T. (1957), *Mental acts: their content and their objects*, Routledge, London.

Gertler, B. (2010), *Self-knowledge*, Routledge, London.

Gombay, A. (2007), *Descartes*. Oxford: Blackwell.

Hacker, P. M. S. (1972), *Insight and Illusion*, Clarendon Press, Oxford.

Hintikka, J.(1962), “Cogito, ergo sum: Inference or Performance?”. In *Philosophical Review*, LXXI: 3–32.

Hooker, M. (1978), *Descartes’s Denial of Mind-Body Identity*, in M. Hooker (ed), *Descartes: Critical and Interpretative Essays*. Baltimore: Johns Hopkins University Press.

Hume D. (1739-40), *A Treatise of Human Nature*, Oxford University Press, Oxford.

James, W. (1890), *The Principles of Psychology*, Cosimo, New York, 2007.

Kaplan, D. (1989), *Demonstratives and Afterthoughts*, in Almog J., Perry J., Wettstein H. (eds) (1989), pp. 481–614.

Kenny, A. (1968), *Descartes: A Study of His Philosophy*. New York: Random House.

Kriegel, U. (2003), *Consciousness as Intransitive Self-Conscious-*

ness: *Two Views and an Argument*, “Canadian Journal of Philosophy”, 33, 103-132.

ID. (2007), *Self-Consciousness*, “Internet Encyclopedia of Philosophy”, <http://www.iep.utm.edu/self-con/>

Mach, E. (1914), *The analysis of sensations*, Open Court, Chicago & London.

Markie, P. (1986), *Descartes’s Gambit*. Ithaca N.Y.: Cornell University Press.

Markie, P. (1992), “The Cogito and its importance”, in J. Cottingham (ed), *The Cambridge companion to Descartes*. Cambridge MA: Cambridge University Press, pp. 144-173.

Matthews, G.B. (1992), *Thought’s Ego in Augustine and Descartes*, Cornell U.P., Ithaca NY.

Noonan, H.W. (1999), *Hume on knowledge*. London: Routledge.

Perry, J. (1977), *Frege on Demonstratives*, “Philosophical Review” 86 (4), 474-497.

ID. (1979), *The Problem of the Essential Indexical*, in Cassam Q. (ed) (1994), pp. 167–83

ID. (1997), *Indexicals and Demonstratives*, in Hale B., Wright C. (eds), *A Companion to the Philosophy of Language*, Blackwell, Oxford, pp. 586-612.

Recanati, F. (2007), *Perspectival Thought. A Plea for (Moderate) Relativism*, Oxford U.P., New York.

ID. (2009), *De re and De se*, “Dialectica”, 63, 249-269.

Sluga, H. (1996), *Whose house is that? Wittgenstein on the self*, in H. Sluga, D. G. Stern (eds.), *The Cambridge Companion to Wittgenstein*, Cambridge U.P., Cambridge MA.

Shoemaker, S. S. (1968), *Self-reference and self-awareness*, “The journal of philosophy”, 65(19), 555-567.

Strawson, P.F. (1966), *The Bounds of Sense. An Essay on Kant’s Critique of Pure Reason*, Methuen, London.

Stroud, B. (1977), *Hume*. London: Routledge & Kegan Paul.

Van Cleve, J. (1983), “Conceivability and the Cartesian argument for dualism”. In *Pacific Philosophical Quarterly* 64: 35-45.

Williams, B. (1978), *Descartes: The Project of Pure Inquiry*. Abingdon: Routledge.

Wilson, M. D. (1978), *Descartes*. London: Routledge.

Wittgenstein, L. (1958), *The Blue and the Brown Books*, Blackwell, Oxford.

Zemach, E. (1985), “De Se and Descartes: A new semantics for indexicals”. In *Nous* 19 (2): 181-204.