

## SECOND MEDITATION

*The nature of the human mind, and how it is better known than the body*

The sceptical hypotheses of the First Meditation had called into question apparently all the beliefs that had hitherto been taken for granted: empirical beliefs in the familiar everyday world of trees and buildings, fires and dressing gowns; beliefs in the entire physical world, 'body, shape, extension, movement and place', including belief in the thinker's own body; beliefs in a *priori* mathematical truths, that were challenged by the Demon hypothesis. Descartes hopes to find just one certainty that will be invulnerable to the sceptical hypotheses, one Archimedean point, and he finds it in the argument: *cogito ergo sum*. This famous formulation of the argument is from the version in Descartes' *Discourse on Method*: '...this truth 'I am thinking, therefore I exist' [is] so solid and secure that the most extravagant suppositions of the skeptics could not overthrow it'. In the *Meditations* he puts the argument like this.

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. 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 let him deceive me as much as he can, he will never bring it about that am nothing so long as I think I am something. So after considering everything very thoroughly, I must finally conclude that the proposition, *I am, I exist*, is necessarily true whenever it is put forward by me or conceived in my mind.

What kind of an argument is this? Is it strictly an argument at all? The traditional formulation, 'I think, therefore I am' looks like an argument in every way: it has a premise ('I think'), a 'therefore' indicating an

inference, and a conclusion ('I am'). On the other hand, Descartes says in reply to the Second Objections:

When we observe that we are thinking beings, this is a sort of primary notion, which is not the conclusion of any syllogism; and, moreover, when somebody says; I am thinking, therefore I am or exist, he is not using a syllogism to deduce his existence from his thought, but recognizing this as something self-evident, in a simple mental intuition.

Descartes' readers have disagreed about whether the cogito is an inference, or a simple 'intuition'. Others have said that the *cogito* is not quite an inference, not quite an intuition, but a 'performance'.<sup>4</sup>

#### *The special status of 'I think' and 'I am'*

Descartes says that there is something special about his belief *that he is thinking*, and his belief *that he exists*. What exactly is special about these thoughts? Descartes says that it is impossible to doubt these beliefs. So what is it about them that makes them immune to doubt? Descartes says in the passage above that 'the proposition *I am, I exist*, is necessarily true whenever it is put forward by me or conceived in my mind'. If these propositions are necessarily true, then that might be why they cannot be doubted. But are they 'necessarily true', as Descartes says? No, or not strictly. There are possible worlds in which Descartes does not exist. Perhaps in those possible worlds, his parents never even met. If Descartes is taken to be the referent for 'I', then in those worlds the proposition 'I, Descartes, exist' is false. The proposition that 'I, Descartes, think' is also false, in those worlds. And something similar will apply no matter who the thinker is. These propositions are not necessarily true, in the usual sense in which philosophers speak of necessary truth.

---

<sup>4</sup> See e.g. Anthony Kenny, *Descartes: A Study of his Philosophy* (Random House, 1968), Ch. 3, and Jaako Hintikka, 'Cogito ergo sum: inference or performance?' in Willis Doney (ed.) *Descartes: A Collection of Critical Essays* (Doubleday 1967), pp. 108-39.

These propositions are contingent. But Descartes is surely right about their special status.

One suggestion is that these propositions have the special character of being *incorrigible*, and *self-verifying*. This suggestion has been made by Bernard Williams (Descartes: *The Project of Pure Enquiry*, Ch. 3). It is assumed in his definitions that we are dealing with contingent propositions: propositions that are not true in all possible worlds. With that assumption (and why does that assumption matter?) he proposes that a proposition *P* is *incorrigible* when it satisfies this description: *if I believe that P, then P*. Here are some candidate examples of propositions that satisfy this description. 'I am in pain.' This is, arguably, incorrigible, since if I really believe I am in pain, I am in pain. I can't be wrong about it. 'That looks red to me.' Here again, if I really believe something looks red to me, then it does look red to me. I am the expert about how things look to me. I might be wrong, of course, on how they really are. But (arguably) I can't be wrong about how they look to me. Now consider the propositions from Descartes' argument. 'I am thinking.' Suppose I believe that I am thinking. It follows that I am thinking. 'I exist'. Suppose I believe that I exist. It follows that I do exist. The propositions 'I think' and 'I exist' both seem to be incorrigible, in Williams' sense.

The propositions are also *self-verifying*. This is a closely related concept, which concerns *assertion* rather than belief. A proposition *P* is self-verifying when it satisfies this description: *if I assert that P, then P*. Here are some candidate examples of propositions that satisfy this description. 'I am speaking'. 'I can speak at least a few words of English'. 'I promise to come to the party'. If I assert (out loud!) that I am speaking, then I am speaking. If I assert that I can speak at least a few words of English, then I can speak at least a few words of English. If I say that I promise to come to the party, then I do promise to come to the party. The latter is an example of what Austin called a *performative* speech act. Some philosophers who see a similarity between this example and the propositions of the *cogito* have developed the 'performative' interpretation of Descartes' argument, mentioned above.

Is there a similarity? Yes, in so far as all are examples of self-verifying propositions. If I assert 'I am thinking', then I am thinking. If I assert 'I exist', then I exist. Contrast these examples with *self-refuting* statements. 'I am absent.' 'I cannot speak any English'. 'I cannot think'. 'I do not exist'. (Can you imagine situations where these propositions might be used in a way that is not self-refuting?)

If this is correct, then there is indeed something special about the status of the propositions of Descartes' argument. Although they are strictly speaking contingent propositions, not necessary ones, they have the special features of being incorrigible and self-verifying. That is why they cannot be doubted.

However, there is a puzzle now. The conclusion Descartes wants to reach is 'I exist'. If this proposition *on its own* has the vital properties of being incorrigible and self-verifying, then why does Descartes bother with his premise, 'I think', and trouble to present the argument as '*I think, therefore I am*'? The answer is not obvious, but here are two suggestions.

The first is that there is one formulation of the argument presented in the Meditations which can be interpreted in just this way. Descartes says, 'the proposition...*I exist*, is necessarily true whenever it is...conceived in my mind.' This could be interpreted exactly in line with Williams' suggestion: 'if the proposition 'I exist' is believed by me (conceived in my mind), then it is true'. In Williams' terminology: the proposition 'I exist' is incorrigible. On this reading, the conclusion 'I exist' is inferred from a thought about one's existence. This still leaves all the other formulations of the argument, however, in which existence seems to be inferred from thoughts about something other than one's existence (25). 'If I convinced myself of something, then I certainly existed.' If there is a deceiver who is deceiving me, 'in that case too I undoubtedly exist'. If I 'thought anything at all, then I certainly existed' (French version). In all these cases the premise that is supposed to yield a conclusion about existence is not a thought *about one's existence*, but rather a thought about e.g. a deceiver.

The second suggestion is that there is indeed a special reason for the premise of the argument being 'cogito', 'I think'. The proposition 'I think' has a special feature that is lacking in the proposition 'I exist'. The proposition 'I think' is *evident* (to use Williams' label). A proposition P is evident to me if it satisfies the following description: *if P, then I believe that P*. Compare this definition to the definition for incorrigibility above. You can see that being evident is roughly the converse of being incorrigible. Incorrigibility says, if you believe it, it's true. Evidence says, if it's true, then you believe it. If something is incorrigible to you, then you are an expert about it, in one way. If you believe it, it's true. If something is evident to you, then you are an expert about it in a different way. If it's true, then you believe it. It doesn't escape your attention.

It may be that Descartes thinks that all propositions about the mind are incorrigible and evident. Incorrigibility says: when I believe something about my mind, I get it right. If I believe some proposition about my mind, that proposition is true. That, on its own, is compatible with there being all kinds of dark corners and alleys of the mind about which I know nothing. But then Evidence adds: I know all there is to know. If some proposition about my mind is true, then I believe it. The poet Gerard Manley Hopkins said, 'O the mind, mind has mountains; cliffs of fall / Frightful, sheer, no-man-fathomed.'<sup>5</sup> Descartes would have disagreed; he seems to think there are no unfathomable depths to the mind. There are no hidden corners or dark alleys. The mind is transparent to itself. I can know about all the operations of my mind.

The proposition 'I think' is evident, in a way that the proposition 'I exist' is (apparently) not. If I think, then I believe that I think. Is it true that if I exist, I believe that I exist? No, or at least it seems not. While I believe that I think, when I am thinking, I do not always believe that I exist, when I am existing. Perhaps I can continue to exist in a dreamless, thoughtless sleep, and surely this is what common sense

---

<sup>5</sup>*Poems and Prose of Gerard Manley Hopkins*, ed. W.H. Gardner, (Penguin 1953), reprinted 1971, p. 61.

supposes. In that case I exist, but do not believe I exist, since I do not believe anything. While 'I exist' is incorrigible, it is not, on the face of it, evident. So Descartes has a reason for choosing to begin his argument with the premise 'I think'.

### *Essence and existence*

There is another reason for beginning with the premise 'I think' in the famous argument of the *cogito*, which brings us to the question not only about the existence of the 'I', but about the nature of that 'I'. Descartes wants to argue that thinking and existence are very closely connected, in the case of a self that thinks. After addressing the question of his existence, the thinker of the meditation will address the question of his own essence or nature. He will argue in the end, 'I am essentially a thing that thinks'. An essential property of a thing is a property which that thing is bound to have, a property that it cannot lack. Perhaps an essential property of a cat is that it is an animal. Perhaps an essential property of a yeti is that it is an animal. Notice, from the last example, that we can talk about the essential properties of things without being committed to the existence of the things. Nevertheless, truths about essence have implications for existence: if a cat exists, it must be an animal; if a yeti exists, it must be an animal. If it were not an animal, it would not be a cat. If it were not an animal, it would not be a yeti.

If Descartes' argument that the 'I' of the Meditations is essentially a thinking thing is successful, then the implication is similar: if I exist, I must be thinking. If Descartes' argument about his essence is correct, he will be able to argue in either direction. I think, therefore I am (*cogito ergo sum*). And, I am, therefore I think. (I am essentially a thinking thing.) This symmetry will be central to Descartes' vision of what it is to be an 'I', a soul, or self, or mind: *I am if and only if I think*. Notice that if this thesis about essence is correct, it will have the consequence that 'I exist' is not only incorrigible, as described above, but also evident: if I exist, then I

will think, and therefore (by the *cogito*) I will believe that I exist. However, this thesis about the essence of the self is yet to be argued for.

You will notice that throughout the *Meditations*, Descartes carefully distinguishes the question of the existence of something, from the question of the essence or nature of that thing. Both kinds of questions concern metaphysics: What exists? What are things like? Sometimes we might know that something exists, without knowing what it is like. You come home in the dark to a house that you expect to find empty, and you hear an ominous rustling inside. Something is there, but you don't know what it's like. Is it a cat? A burglar? A friend? In such a case you might say, in Descartes' terms, that you know of the existence of something, but you don't yet know its essence, its nature. Sometimes it might be the reverse. You know what Santa Claus is like, you know what his nature is: an old man, white-bearded, red-suited, with a generous disposition and a hearty laugh. (Perhaps not all of these properties belong to his essence. Could there be a young Santa? Or a beardless one? Or a female one?) You know, more or less, what Santa is like: but does he exist? In this case you know what his nature is before you settle the question of whether he exists. The sceptical arguments of the First Meditation have, in general, left the meditator in ignorance about the *existence* of things. The meditator knows what trees, fires and dressing gowns would be like: but he is not sure whether there are any. He knows what his body would be like, if it existed (something with hands, extended in space, etc.), but he is not sure whether it does exist.

This pattern is typical of the *Meditations*. The meditator typically begins by answering some question about essence, and then raises the question about its existence: he will begin by describing the essence of some kind of thing, whether bodies, or shapes, or God, and thereafter raise the question of whether that thing in fact exists. Descartes assumes that the essence of a thing can generally be known before one knows whether the thing exists, because the essential properties of a thing are implied by the idea or concept of that thing. For example, the idea of a triangle implies the essential properties of a triangle: a closed three-sided

figure. That tells you something about the concept of a triangle. And it tells you something about the world: if you come across an existing triangle, it will have three sides. The idea or concept of a yeti implies that being an animal is an essential property of a yet. That tells you something about the concept of a yeti. And it tells you something about the world: if you come across an existing yeti, it will be an animal. You can have an idea or concept of a thing prior to knowing whether the thing exists: so in many cases you can know the essence of something before knowing whether it exists.

The grand exception to this general pattern in the *Meditations* is, of course, knowledge of oneself. In the argument of the *cogito*, the thinker concludes, 'I exist'. It is only *after* establishing this conclusion about his existence that he raises the question: what am I? What is my nature? What is my essence? Immediately after the conclusion of the *cogito*, Descartes says: 'I do not yet have a sufficient understanding of what this 'I' is, that necessarily exists'. I know *that I exist*, but I do not yet know *what I am*. I know of my existence, but I do not yet know of my essence.

The rest of the Second Meditation is devoted to arguing that the essence of the self, or 'I' (whose existence has been proved in the *cogito*) is *to think*. Notice that the subtitle of this Meditation is basically devoted to this issue about the nature of the self or mind: 'The nature of the human mind, and how it is better known than the body'. Notice the mention of 'body' in the title. The meditator will address an important question about the essence of *matter*, or body, whose existence is still entirely in doubt. But the purpose there too will be to establish a thesis about the mind: that it is better known than the body.

#### *Essence of the self, or mind*

The argument of the *cogito* concludes 'I exist': but who or what is it that exists? Not a human body. Not a soul in the traditional Aristotelian sense. Aristotle had identified the soul with certain capacities that living things possess: capacities of nutrition, reproduction, locomotion, perception, and thought. On the Aristotelian



account, all living things have souls: plants have the first two capacities, non-human animals have the first four, and human beings have all five. Descartes considers four of these capacities (coyly omitting reproduction), and argues that none but the last capacity, thought, is essential to the nature of the soul.

What about the attributes I assigned to the soul? Nutrition or movement? Since now I do not have a body, these are mere fabrications. Sense-perception? This surely does not occur without a body, and besides, when asleep I have appeared to perceive through the senses many things which I afterwards realized I did not perceive through the senses at all. Thinking? At last I have discovered it - thought; this alone is inseparable from me. I am, I exist - that is certain. But for how long? For as long as I am thinking. For it could be that were I totally to cease from thinking, I should totally cease to exist.

He somewhat overstates his case here, to emphasize the point: the meditator is not *denying* the proposition 'I have a body', but rather *refusing to assent* to it, since the arguments of the First Meditation show that it is dubitable. I can doubt that I have a body: so I can doubt that I have any of the bodily capacities described on the Aristotelian picture, whether of nutrition, locomotion, (reproduction), or perception, in so far as that involves bodily sense organs. (Notice that in so far as perception has a mental aspect, Descartes will treat it as a mode of thinking.) All capacities other than thought are vulnerable to the sceptical arguments of the First Meditation. Descartes concludes that his essence is to think. *Sum res cogitans* : 'I am a thing that thinks'. Notice that Descartes appears to believe he has established not only 'I think'; not only 'I am a thinking thing'; not only 'thought is a property essential to me'; but the strong conclusion that 'thought is the *only* property essential to me'.

### *Essence of body*

We have knowledge of the self: but surely, so the naive view runs, our knowledge of bodies, through the senses, is still more distinct? Descartes considers our knowledge of a particular body: a piece of wax. Or rather, since he has not yet countered the sceptical arguments of Meditation I, he is considering the concept of a particular piece of matter, without committing himself to its existence. He is conducting a kind of thought experiment. Suppose that I were to have knowledge about a material thing. What would its essence be? And how would I know it? The wax is white, scented, hard, cold: these seem to be properties that enable me to understand it distinctly. All those properties disappear when it is placed by the fire: but the thing still continues to exist. 'What was it in the wax that I understood with such distinctness?' Not the fragrance, hardness, coldness, but merely something 'extended, flexible and changeable'. This is a variant of a 'think away' argument, to discover the essential properties of something. (Think away Santa's white beard. Could he still be Santa? If so, then the white beard is not essential to him.) Unfortunately it is not quite clear what essence Descartes is trying to discover: it is not quite clear whether he is asking a question about the essential properties of wax in general, or a particular lump of wax, or of matter in general—questions which would all have different answers (what might they be?). Here it will be assumed that Descartes intends to discover the essence of matter in general.

Descartes reaches a conclusion about the essence of matter. He concludes that the concept of 'body' is the concept of something essentially *extended*, with shape and size, capacity for change of shape and size, and that is all. This anticipates the mind/body dualism, and the doctrine of primary and secondary qualities, developed more fully in the later Meditations.

One question about matter concerns its essence: what would matter be like, if it existed? Another concerns our knowledge: how would we have knowledge of matter, if it existed? Descartes reaches the apparently radical conclusion that bodies, or rather the essential properties of bodies, are known not by mere sense

perception, or imagination, but the *intellect*: perception always involves judgment. This applies to the sensory perception of all material bodies. It applies to the sensory perception I would have of the wax, if it were to exist. And it also applies to perception of the most mundane things: the people I seem to see outside the window. Even if perception were veridical (which the First Meditation gives us reason to suspect), perception would not yield acquaintance with the people themselves, obscured by hats and coats. To judge that they are men is to go beyond perception would tell us: it is to use one's intellect.

Descartes concludes this Meditation with some more discoveries about the self. Knowledge of the self, or mind, is more distinct and certain than knowledge of body. The knowledge of the self given by the *cogito* argument is prior to knowledge of body, and immune to sceptical worries about body. Moreover, every judgment about body helps me to know myself better. 'Every consideration whatsoever which contributes to my perception of ...body, cannot but establish more effectively the nature of my own mind'. This follows from the thesis about the transparency of the mind to itself: in Williams' terms, the thesis about the *evidence* of propositions about the mind. If I judge that there are men below in coats and hats, then I know that *I judge* that there are men below in coats and hats, so I know something not only about them, but about *myself*. The more I learn about anything else, the more I learn about me.

### *Conclusion*

The method of doubt in the First Meditation appeared to threaten all knowledge, but in the Second Meditation the thinker finds something that cannot be doubted. Having tipped out the barrel, the thinker finds one apple that is sound. Having demolished the building, he has discovered a piece of timber that is firm, and that can (he hopes) form the foundation for rebuilding the edifice. Moreover, if his arguments have succeeded, he has discovered the essence of mind, which is to think; and the essence of matter, which is to be extended. And

he has discovered that, contrary to common sense, the mind is more knowable than the material world.

*Some questions to consider about Meditation II*

(1) How plausible is it that propositions concerning your own mental states are incorrigible to you? Isn't it possible to make mistakes about our own beliefs and desires? I might falsely believe that I like the taste of beer, when really I hate it, but pretend to everyone *including myself* that I like it, so I can be one of the crowd at the pub. Is that a possibility? Many people, and many philosophers, think that self-deception is possible. For example, if someone 'turns a blind eye' to the lipstick on her husband's collar, she somehow pretends to herself that he is faithful. In such a case she will have beliefs about her mental states that are not incorrigible: she may believe that she believes he is faithful; but in fact she believes he might not be. Let R be the proposition 'I believe he is faithful'. In this example, she believes that R; but in fact not-R is the case. She is wrong about what she believes. R is not *incorrigible*.

(2) How plausible is it that propositions concerning your own mental states are evident to you? If self-deception is possible, then that is relevant to the evidence issue too. There may be truths about one's mental life about which one is not aware. The woman in the previous example may believe 'deep down' that her husband is having an affair: but that belief is not transparent to her. She does not believe that she believes it. Let S be the proposition 'I believe he is having an affair'. In this example, S is the case; but she does not believe that S. She has a belief of which she is not aware. S is not *evident*. More generally, if there are unconscious mental processes, as many (most famously Freud) have argued, does that undermine Descartes' view?