

## DESCARTES' DUALISM VERSUS BEHAVIOURISM

Andrea Christofidou\*  
Worcester College, University of Oxford

How little do we know of the business of the earth,  
not to speak of the universe; of time,  
not to speak of eternity ...  
Edward Thomas *The South Country*

*Abstract:* My analysis straddles Descartes' metaphysics and some parts of contemporary philosophy, especially regarding consciousness, and aims to show that once our understanding is freed from philosophical habits that affect current debates, Descartes' views offer an opportunity to draw important insights. Primarily, I examine Descartes' mind-body dualism and contrast it with behaviourism, particularly with philosophical behaviourism, focusing on Gilbert Ryle's dispositional behaviourism and his attacks on Descartes' dualism. The discussion takes the form of *Objections and Replies*, presenting the two thinkers in some sort of dialogue with one another. This brings out clearly who is distorting our ordinary language, violating the logical geography of concepts, committing a category mistake, and systematically misleading us. Ryle's two well-known accusations – the category mistake, and the dogma of the ghost in the machine – are turned, by a *reductio ad absurdum*, against his own commitments, leading to an evaluation of his highly paradoxical view, and showing how it collapses in on itself. The closing parts touch upon, but do not pursue, some fundamental concerns about personhood and the self, the metaphysics of mind, freedom, and moral significance, and raise the question of what our deepest concerns and responsibility in the twenty-first century must be.

*Key words:* Descartes, dualism, mind, self-consciousness, Ryle, dispositions

My primary undertaking is to elucidate and examine Descartes' mind-body dualism,<sup>1</sup> and to contrast it with behaviourism.<sup>2</sup> Given the varieties of behaviourism, which cannot be dealt with in a single paper, my concern is not with all the different forms of behaviourism that have been put forward in the last one hundred years, including those that figure in cognitive science, or in scientific psychology,<sup>3</sup> but with philosophical behaviourism and its opposition to Descartes' dualism. Even among philosophical behaviourism there is divergence among its various proponents. Given that the central topic concerns Descartes' dualism, I shall focus on Ryle's (1949) philosophical dispositional

---

\* My thanks are first and foremost to José E. Burgos for inviting me to address the debate between dualism and behaviourism, for his various comments and suggestions, and for his sustained encouragement. My thanks also go to Peter J. King for numerous discussions, for his critical and constructive suggestions, and for all his support and encouragement throughout.

<sup>1</sup>All references to Descartes' works are given in the text, using the standard citations.

<sup>2</sup>My use of the term 'behaviourism' in the singular is not meant to suggest that there is one kind of behaviourism; this is made clear in what follows.

<sup>3</sup>This has been explored recently by, among others (Harzem, 2004, pp. 5–12; see also Burgos & Killeen, 2018). Not even O'Donohue and Kitchener's (1999) handbook can cover all the different varieties of behaviourism and the disagreements among its practitioners, as its editors admit.

behaviourism because it is the most direct attack on Descartes' dualism, and it is the best known. My analysis is not primarily historical, but straddles Descartes' metaphysics and (some parts of) contemporary philosophy; it aims to show not only a relevant continuity between his conception and our contemporary debates, but also to reveal that, once our understanding is freed from the philosophical habits that affect contemporary debates, his views offer an opportunity to draw important insights.

In light of this, a preliminary consideration of some contemporary debates is in order. This I address in the first section. In the second section, I elucidate and discuss Descartes' dualism. In the third section, I offer a brief characterisation of behaviourism and philosophical behaviourism. The fourth section is where Ryle's objections to Descartes' dualism are presented and discussed; the discussion takes the form of '*Objections and Replies*', presenting the two thinkers in some sort of dialogue with one another. This will bring out clearly which one is distorting our ordinary language, violating the logical geography of concepts, committing a category mistake, and systematically misleading. It will also enable us to consider some fundamental concerns about personhood and the self, about the metaphysics of mind, of freedom, of moral significance. And it will enable us to appreciate the immense difference between Descartes and Ryle in their approach to proper philosophical enquiry and analysis.

The fifth section assesses Ryle's two main accusations – the category mistake, and the dogma of the ghost in the machine – against his own commitments, juxtaposed with Descartes' commitments. The sixth section offers an overall evaluation of Ryle's highly paradoxical view and shows how it collapses on itself. The conclusion, in the form of an Epilogue, considers the question of philosophical habits perpetuated from one generation of thinkers to another, and raises the question of what our deepest concerns and responsibility in the twenty-first century – morally, epistemically, metaphysically – must be.

### Some Preliminary Considerations

The last part of the twentieth century began to see the waning of physicalism that had been dominant for more than half a century and fiercely defended by its proponents. The primary reason for its waning is that physicalism cannot account for the mind and its distinctive essential nature: it cannot explain either its conative or its cognitive acts or states — it cannot accommodate or explain consciousness, nor can it accommodate or explain self-consciousness, reason, freedom, and the self who is presupposition of, and inseparable from them. "We cannot distance ourselves from [them]. That is Descartes' insight." (Nagel, 2012, pp. 81–82).

In fact, physicalism faces a dilemma within its own thesis (see Hempel, 1980, pp. 193–200): how to understand what it is to be physical, or how the term 'physical' is to be defined. First, defining the physical as non-mental is simply uninformative, despite Papineau's (2001) claim that this "is just what we need for philosophical purposes" (p. 12; see **An Evaluation of the Argument** in the second section); defining it in terms of whatever our current physics tells us there is renders it not only circular, but is little more than an appeal to authority — "though oddly, to an authority which is essentially mutable and incomplete" (King, 2006, p. 63). It's the other side of the coin, which Descartes questioned and abandoned: appeal to the authority of theology. Secondly, defining, or attempting to offer an understanding of the physical by appealing to whatever a complete future physics would include, renders physicalism vacuous, with no determinate content, and trivial, with no explanatory power. Nor does it follow that any future physics would bear a likeness to our current conceptions, or that the completeness of physics would be the totality of truths, unless the physicalist conflates the totality of truths with the physicalist conception (Christofidou, 2007; 2013, pp. 226–235).

As a result, mind–body dualism underwent what seemed a renaissance during that period and into the turn of the twenty-first century and was considered a leading alternative to physicalism. But, as is usually the case, various other kinds of theory in the metaphysics of mind besides dualism garnered interest, perhaps most surprisingly a resurgence of behaviourism (to which I shall return in Section 3), but also a number of theories purporting to be non-physicalist while rejecting Descartes' dualism.

Philosophers, such as Chalmers (1996), state that it's widely agreed that consciousness arises from a physical basis, but that we have no explanation of why and how it so arises. It seems physically and scientifically unreasonable that it should, and yet it does, the proponents of the thesis state. The central question of the thesis, dubbed 'the hard problem of consciousness,' is: 'why and how do physical processes in the brain give rise to consciousness?' Presumably, those who accept and defend the thesis that consciousness *does* arise from physical processes in the brain, must have good reasons for it beyond a mere aversion to dualism, though it's hard to see how one could have good reasons to hold a thesis that didn't include at least some explanatory account of how the thesis might be true. In the absence of any such account, we seem to be faced either with a vicious circle (we state that *a* does give rise to *b*, then we look for an explanatory account of how *a* might give rise to *b*, which then justifies the statement that *a* does give rise to *b*), or with the initial statement widely accepted for no good reasons.<sup>4</sup>

The proponents of the thesis argue that the hard problem of consciousness is the number-one unanswered scientific challenge of our times. I suspect that many scientists in areas unrelated to consciousness would take issue with this, citing for example how little we know of the planet Earth, of dark matter, of the universe as a whole (as the quotation at the head of the paper states), though it's true that the problem is hard for science; but that's because it's a metaphysical problem, and not amenable to scientific investigation.

One suggested answer to the hard problem of consciousness involves the revival of a thesis known as panpsychism (which can be variously interpreted and understood). According to the recent revival of the thesis, consciousness is a primitive fundamental component of reality in the same way that space and time are.<sup>5</sup> I am drawn to and welcome the thesis that consciousness or, more accurately, the mind (because consciousness doesn't float about without a mind; it's ontologically dependent on the mind) is a fundamental, primitive – that is, unanalysable – irreducible entity. What I find difficult to reconcile regarding the hard problem of consciousness are the two claims put forward by proponents of the thesis. On the one hand: 'physical processes in the brain give rise to consciousness;' the problem is, why and how? On the other: 'consciousness is a fundamental component of reality;' the task is to discover the fundamental principles that govern consciousness, and the fundamental laws that connect consciousness to physical processes in the brain. But if consciousness *does* arise from the physical processes of the brain, how can it be a fundamental component of reality? If it is a fundamental component, how can it arise from the physical, or from anything else for that matter? What is fundamental cannot arise from or be dependent on something else.

More recently, a thesis referred to as neo-Russellian monism is put forward again as a non-physicalist thesis<sup>6</sup> while also defining its position against Descartes' dualism. This opposition, however, is based on a number of misunderstandings of Descartes' modal metaphysics: for example, neo-Russellian monism claims that for Descartes the mind is separate from the body (and thus the brain). But as we shall see, Descartes argues that the mind *can* exist separately from body, *not* that it *does* exist separately. Neo-Russellians argue that physics and other special sciences tell us surprisingly little about

---

<sup>4</sup>It might be argued that the claim, 'physical processes in the brain give rise to consciousness', is an hypothesis. However, unlike in sciences where an hypothesis can be tested empirically, this hypothesis cannot be, and in the absence of any good reasons or an argument, it's not derivable *a priori*. Neuroscience might be empirical, but sheds no light on consciousness; at best it might show that there are causal connections between brain processes and conscious experiences. But that's neither illuminating nor new. Descartes defends a causal interaction between the mind and the brain, or the human body.

<sup>5</sup>This assumes that space and time are widely accepted to be real and fundamental components of the universe, but such controversies don't affect the present case.

<sup>6</sup>As to what the physical is, the question is considered to be terminological; the physical is non-mental — which is, as we have seen, utterly uninformative. See, among others, Goff (2017a, p. 23).

what the brain *is*, its intrinsic nature, even if it tells us what the brain *does* — its extrinsic properties.<sup>7</sup> The brain is a physical real particular that has shape and size, breadth and depth, colour, smell, and taste, synapses and neurons, but all these, including its structural and relational properties, are extrinsic to its nature. As neo-Russellians state: “grasping the essence of a property does not entail grasping the essence of a property that grounds it” (Goff, 2017a, p. 198). What we know is that “brains [...] have consciousness-involving nature” — the intrinsic nature of brains; put differently, “the irreducible reality of the mind just is the intrinsic nature of the brain” (Goff, 2017b).

But how do we know this? Or is this again an hypothesis? It’s also odd, to say the least, that the intrinsic nature of something, what makes it *what it is*, can be radically different, that is, distinct in *kind*, from its extrinsic properties and from *what it does*. It’s equivalent to saying that the intrinsic nature of fire is ice, despite its extrinsic properties being heat and flames – and what it does (burning wood, melting wax, etc.) – are radically different from its intrinsic nature. Even if we accept that the extrinsic properties of the brain are not *constitutive* of its intrinsic nature, they must surely be *consequences* of that nature, unless one accepts Leibnizian well-founded phenomena: that physical phenomena, such as brains, are not themselves real, but are only the way we perceive what is in fact a combination of the true and real immaterial monads or substances. For neo-Russellians, however, the brain is a real physical entity. Moreover, accepting Leibnizian well-founded phenomena, neo-Russellians must accept that there are minds which are ontologically distinct complete entities, that is, mental substances. This would evidently contradict the core of their thesis that minds are not distinct complete entities. I shall leave these and other concerns for another paper.

What’s pertinent for present purposes is that neo-Russellian monism is opposed to Descartes’ dualism, and instead claims to be a general thesis for unifying matter and mind. If the mind is distinct in kind from matter, requiring a distinctive metaphysical account as neo-Russellians seem to be saying (Goff 2017a, pp.145–147), how does the unity of matter and mind avoid Descartes’ dualism? This is another misunderstanding by neo-Russellians, since Descartes argues for, and defends a true substantial union between mind and the human body, a union “ordained by nature” (see his *Optics* AT VI 130; I shall come back to this later.). Neo-Russellians’ response would be that they’re avoiding Descartes’ dualism because they’re not committed to two entities: mind and brain (or body). It seems to follow from this that for neo-Russellians the term ‘mind’ refers to an incomplete entity, to consciousness, which must be a property. The question is: in what sense are properties fundamental? Could a property be ontologically fundamental if it has no self-sufficiency, that is, if it’s not a complete entity but a dependent entity? If so, is consciousness only explanatorily fundamental? These are not problems that can be addressed here, but it seems that the status of consciousness cannot be on a par with that of the brain, which is a complete entity and hence ontologically fundamental.

Neo-Russellians’ rejection of Descartes’ dualism is not based on an argument but on an assumption, often admitted with startling frankness: “Assuming the falsity of substance dualism, we know how to colour in one bit of [the world]: the brains of organisms are coloured in with consciousness” (Goff, 2017a, p. 171). It’s not only unclear what the force of this claim is, but assuming the falsity of a thesis is hardly equivalent to proving its falsity. Surely, in philosophy what are required are cogent arguments.

What unifies all these positions in their rejection of Descartes’ dualism is their inveterate adherence to the physical causal closure: an unfounded and unargued dogma, being neither based on empirical evidence nor derivable *a priori*. It’s thus a mistake to take it as a guide to ontology or metaphysics, yet it has dominated philosophy for more than seventy years and been glorified as a principle. For example, Chalmers (1996) states: on “the dualist view I advocate [...the] causal closure of the physical is preserved” (p. 170). But if the causal closure of the physical is preserved (and why should this matter?), then mustn’t the mind or consciousness be idle, an epiphenomenon without any

---

<sup>7</sup>Neo-Russellians in fact argue that physics tells us very little or nothing about the intrinsic nature of the physical world.

efficacy to affect or be affected by anything? The question then is: why postulate, or even talk about consciousness? Or, does consciousness exist simply in words, not in reality?

Reductionist and non-reductionist physicalists, neuronists, non-physicalists, and neo-Russellian monists argue that their goal is naturalistic, a naturalist approach to reality, as against Descartes' dualism. A putative ineradicable tension between dualism on the one hand, and all these positions on the other, seems to hinge at least on this notion of naturalness. The term 'natural', and its cognates 'naturalistic' and 'naturalism' is one of the terms most misused in philosophical debates in the last hundred years or so. But as Chalmers (1996) states: "There is no *a priori* principle that says that all natural laws will be physical laws; to deny materialism is not to deny naturalism." (p. 170). A more recent objection to Descartes' dualism claims that it "fails to integrate consciousness sufficiently in the natural causal order" (Coleman & Alter, in press). It's not clear what 'sufficiently' is meant to convey here, but the objection demonstrates both a misuse, a misappropriation of the notion of natural, and a lack of understanding of Descartes' dualism. When Descartes refers to the true nature of mind and the true nature of body, he's not referring to anything supernatural, or non-natural. His dualism is natural through and through. When he refers to the natural light of reason (which is not an arbitrary assumption but a precondition of self-consciousness) he refers neither to physical nor to supernatural or divine illumination — he rejects all these. Moreover, as we shall see in the second section, causality is a metaphysically basic category, and neutral with regard to the nature of causal relata. What makes it the case that something is natural or non-natural, what is actual, possible or necessary, and the relations between them, is a question of metaphysics not of epistemology, or physicalism, or scientific psychology, or neo-dispositional/power behaviourism. It may be that everything physical is natural, but it's a philosophical mistake to claim that everything natural is physical, or scientific, or against dualism. This is scientism, not science.

### **Descartes' Argument for Dualism or the Real Distinction between Mind and Body**

Although it is well known, even outside philosophy, that Descartes is a dualist about mind and body, his views are widely misunderstood — not only outside philosophy but even more within philosophy, as are his reasons for dualism and the argument he gives for his position. It's therefore worth at the outset clearing up some of the misconceptions.

First, the relevance of religion. One of Descartes' great achievements (which to a large extent changed the course of philosophy and science) was the abandoning of Scholasticism and the divorce of philosophy from theology. Aside from the fact of his insistence that theology and philosophy be kept separate, most of his critics at the time who argued for a physicalist position were themselves religious (often priests and monks). Gassendi acutely comments on "the highly intelligent and successful way in which [Descartes] attempted to extend the boundaries of the sciences and explain matters which have remained highly obscure in all previous ages" (*Fifth Set of Objections* AT VII.256). No doubt, all this can be subject to philosophical debate, but Descartes' defence of dualism is not a case of religion or theology. In fact, the Church's teaching was centrally that life after death involves the resurrection of the body, and dualism isn't needed for that.

It's interesting that Descartes explains in the *Synopsis* to the *Meditations*, not only that his argument for dualism doesn't entail the immortality of the soul, but that he doesn't have any argument for it; this is because "the premises which lead to the conclusion that the soul is immortal depend on an account of the whole of physics" (AT VII.13) — something the physicalists perhaps need to rethink (see also fourth section).

There are Scholastic, Augustinian, and classical antecedents to Descartes' philosophical concerns; he certainly didn't write in a vacuum. My aim, however, is to attend carefully to Descartes' notions and arguments; as he says: "I wish to point out here that I am paying no attention to the way these terms

have lately been used in the Schools [...] when my own views are profoundly different” (*Rule Three* AT X.369). What’s fresh and radical is often, as he says, the use he makes of the notions. From a Fregean perspective, their new *sense* is *designatum*-involving; it’s not only sense (*Sinn*), a mode of presentation, but also reference (*Bedeutung*). For example, Descartes’ use (in his various *Replies to Objections*, the *Principles*, the *Correspondence*, and elsewhere) of what is translated as ‘soul’ is to be understood neither in a theological sense as spiritual or supernatural,<sup>8</sup> nor in Aristotle’s sense as *anima*, but as mind (*mens*): “I use the term ‘mind’ [*mens*] rather than ‘soul’ [*anima*], since the word ‘*anima*’ is ambiguous and is often applied to something corporeal [as in Aristotle’s use of *psyche*]” (AT VII.161 *Definition VI*). “*Anima* in good Latin signifies *air*, or *breath*; it is in a transferred sense, I think, that it means *mind*. That is why I said that it is ‘often taken for a corporeal thing’” (Letter to Mersenne 21 April 1641 AT III.362; CSMK:180). Aristotle distinguishes between *psyche* and *nous*: thinking mind or intellect (*Peri Psyches*, book III; see Christofidou, 2009, pp. 134–62).<sup>9</sup> Descartes rejects Aristotle’s real distinction between the sensory soul (*psyche*) and the intellectual mind (*nous*), and argues: “I consider the mind not as a part of the soul [*anima*] but as a thinking soul [mind, *mens*] in its entirety” (*Fifth Set of Replies* AT VII.356; see also **An Evaluation of the Argument** below).

Descartes nevertheless includes God in his philosophical discussion; this is often little more than a way of expressing non-religious notions and has no theological commitments (as when a modern atheist talks of seeing space–time ‘from a God’s-eye point of view’), but sometimes it plays a more central role, a way of expressing metaphysical notions such as fundamentality or grounding, totality or completeness. In such cases, Descartes only includes what he can argue for, and makes use of it only where necessary. Nothing that he argues about his dualism depends upon or is founded in religious belief.

Secondly, the separateness of mind and body. Descartes does not argue that minds and bodies are separate; rather, as we shall see (in the next three subsections), he argues that they are *separable*, and thus really distinct. It’s important to attend carefully to his modal metaphysics — that is, to his use of terms such as ‘can’ and ‘could’ and notions such as possibility, necessity, and so on.

Thirdly, he doesn’t argue for dualism in the *Second Meditation*, but has to wait until the last, the *Sixth Meditation*, before he’s in a position to set out the argument (as I shall demonstrate below). Finally, Descartes offers one argument for the *Real Distinction between Mind and Body*; some commentators claim to find more, citing, for example, Descartes’ discussion of the divisibility of the body and the indivisibility of mind — but careful reading shows us that Descartes is here offering a characterisation of the distinction once he has demonstrated it, not arguing for it again (AT VII.85–86). Such a characterisation seems to be directed towards those who might be unable to follow the argument; he’s therefore appealing to something that they’re familiar with, namely, the Scholastics’ thesis that — unlike minds, which were regarded as indivisible — matter was inherently composite and divisible. Descartes himself rejects the Scholastics’ thesis because what he considers as divisible is “one piece of matter” (*Principles* II.25), not material substance. ‘Divisible’ means capable of being divided, at least “in my thought,” (AT VII.86) and for Descartes it’s not a mereological thesis but a commitment to the continuous nature of a body (he also holds a non-mereological thesis for time and space).

In the sub-sections that follow, I set out the main philosophical categories,<sup>10</sup> principles, and notions essential for understanding Descartes’ argument and his conclusion, as well as a number of principles and various distinctions central to the argument, before going on to present and examine the argument itself.

---

<sup>8</sup>Those who claim that for Descartes the mind is spiritual or supernatural demonstrate a certain ignorance of his philosophy.

<sup>9</sup>Aristotle exempts *nous* from hylomorphism.

<sup>10</sup>Categories are basic and unanalysable. ‘Category’ here denotes a principal distinction of what there is, ordered in terms of completeness and independence/dependence relations. See also **Category Mistake** in the fourth section below.

## Substance and Attribute

Descartes' position is known as substance dualism, as opposed to property dualism or conceptual dualism. What, then, is a substance? The notion of substance has a long history stretching back to the pre-Socratic philosophers, who introduced it in response to a serious problem with which they were grappling: the problem of change and identity. If what we consider as things are a collection of properties, it would be very difficult to account for what unites those properties, or to know what it is that individuates one collection of properties from another. Furthermore, the fact that properties can and do change implies that nothing would remain the same across time. Consequently, the principles of unity, individuation, and identity collapse. But without such principles, how are we to make sense of and understand the world? In fact, without true unity the notions of change, multiplicity, and diversity would make no sense. There must, therefore, be something that *has* properties (a subject of properties), and which remains the same through change, multiplicity, diversity. The notion of substance served that primary and indispensable role.

In terms of the category of substance, Descartes distinguishes between infinite and finite substance. Infinite substance is supremely perfect and independent: "By *substance* we can understand nothing other than a thing which exists in such a way as to depend on no other thing for its existence. And there is *only one* substance which can be understood to depend on no other thing whatsoever, namely God [the infinite substance]" (*Principles* I.51; emphasis added). This, says Descartes, is the only substance in the primary, strict sense. "In the case of all other substances, we perceive that they can exist only with the help of God's concurrence" (*Principles* I.51). In this context, 'perfection' denotes completeness (from *perfectionem*<sup>11</sup>); an infinite entity would be supremely complete or perfect, whereas a finite entity can only be complete or perfect of its kind. The infinite substance has an infinity of attributes. Each finite substance has one principal attribute.

The category of finite substance consists of two kinds of substance: thinking substance and corporeal substance. Each finite substance *can* exist independently of other finite substances. Anyone who claims that in Descartes' system, mind and body, *do exist* independently, shows ignorance of *Principles* I.51–52, the *Fourth Set of Replies*: "the notion of a [finite] *substance* is just this — that it can exist by itself" (AT VII.226), among other works. Substance, infinite or finite, has attributes (see below) but no parts; it's a complete entity, a true and real unity. Unlike the unity of a building, a machine, a computer, or a robot, the unity of substance is grounded in reality; it's neither imposed nor constructed by us. Substances, corporeal or thinking, and their respective principal attributes, are indivisible: unity, simplicity or indivisibility, and inseparability are constitutive of *what it is to be* a substance.

In the *Definitions*<sup>12</sup> (appended in the *Second Set of Replies* (AT VII.160–2) Descartes explains: "The substance in which thought immediately resides is called *mind*" (*Def.* VI). "The substance which is the immediate subject of local extension and of the accidents [or properties] which presuppose extension, such as shape, position, local motion and so on, is called *body*" (*Def.* VII). Corporeal substance is body — not *a* body. Sometimes, however, Descartes uses the *ordinary* sense, prevalent at the time, of 'thing or substance' (*res sive substantia*) to refer to *a* body, such as the human body, or a

---

<sup>11</sup>'Perfections' (plural) can mean attributes, though not properties. The more perfections something has, the more reality or perfection it has. Degrees of reality or perfection denote degrees of power and independence (see Christofidou, 2019).

<sup>12</sup>Note that these are "real definitions" — that is, definitions founded in reality, not merely linguistic or conventional definitions.

stone (though a human body differs immensely from a stone; see especially **Descartes' second rebuttal** in the fifth section below).

Attributes are real but incomplete; they're neither ontologically, nor metaphysically, nor explanatorily independent of the substance to which they pertain. Attributes are not properties but constitute the nature of substance — what a substance is, not what we decide to call it. We can think of an attribute as the ability of a substance (by its own nature) to act, to bring about an effect, but also to have modes of a certain kind. Modes are the ways by which the principal attribute of a finite substance is made known in a determinate way (infinite substance has no modes). Thus, to say that a substance has the attribute of thought is to say that it's the sort of substance that can have modes under its principal attribute, which are its various acts of thinking, of willing, of doubting, etc. Similarly, to say that a substance has the attribute of extension is to say that it's the sort of substance that can have modes under its principal attribute, that is, bodies such as stones, mountains, trees, etc.

### **Kinds of distinction**

There are three main kinds of distinction: modal, conceptual, and real. Descartes' dualism is concerned with a *real distinction*, which I shall discuss in 2.3. Here, I discuss modal and conceptual distinctions, and also some of the concepts and misunderstandings associated with the three kinds of distinction in general.

*A Modal Distinction:* the distinction between substance and its modes is a modal distinction. There cannot be modes without substance, but the relation is asymmetrical, since it's logically possible that there be substance without modes. It's also the case that there cannot be attributes without substances, since attributes are not complete entities and hence cannot be independent; this is what Descartes means when he says: "we know by the natural light [of reason] that a real attribute cannot belong to nothing" (*Def. V*). However, it's also the case that a substance could not be *what it is* without its principal attribute; the "only idea we have of a substance itself, in the strict sense, is that it is the thing in which whatever we perceive [clearly and distinctly] exists" (*Def. V*; 'in' denotes dependence, of course, not a spatial relation; substances are *not* containers).

*A Conceptual Distinction:* attributes and substances cannot be separated, and yet we have drawn a distinction between them. This is a conceptual distinction: a distinction drawn by reason, not a *real distinction*. We cannot clearly and distinctly conceive substance independently of its attribute, and we cannot clearly and distinctly conceive attributes independently of substance. This impossibility is not psychological or epistemological, but logical or metaphysical. Without an attribute, substance would not simply be unintelligible, there would be no substance, and *vice versa* (*Principles* I.62). It follows that substance and attribute are inseparable. The substance-attribute conception invokes the independence conception of substance such that, we can reason to an understanding of finite substance as "the kind of thing that can exist without any other created thing" (Letter to Hyperaspistes August 1641 AT III.429; CSMK: 193); the infinite substance is supremely independent.

*Conceiving versus Imagining.* Descartes is concerned with *conceiving*, not with *imagining*; the latter is too crude and limited to be used as a philosophical tool (unfortunately the two are often used interchangeably by modern writers, both in the *Philosophy of Mind* and in *Metaphysics*). As he argues, we can imagine a triangle (that is, we can have something like an image of a triangle in our mind), but imagination cannot reveal the true nature of an entity; only conception or clear and distinct understanding can demonstrate its nature, *what it is to be* a triangle. The distinction can be brought out more clearly when we consider a polygon such as a *chiliagon* (a thousand-sided figure), which we can clearly and distinctly conceive and understand, though we cannot imagine it (AT VII.72). Furthermore, what we *can* imagine, however clear it might be, cannot be distinct, because imagination involves both reason and the senses, and therefore is mixed, not distinct, and thus cannot lead to *scientia*. Descartes



defends our reason's capacity to discover what is objectively true and real.<sup>13</sup> Any objections to Descartes' dualism based on what can be imagined are thus irrelevant to and ineffective against his argument.

In fact, for Descartes, knowledge (*scientia*) and understanding are not based simply on what we can or cannot conceive. As we shall see below, it's the nature of things that constrains our clear and distinct conception and understanding, not the other way around. In other words, our understanding of the nature of the mind is not based on imagining and question-begging thought-experiments about mindless zombies, or lonely bodiless ghosts. Nor is it based on introspection, which has become dominant in recent theories in the Philosophy of Mind and Metaphysics. The arguments in some contemporary theories of the mind proceed from phenomenology – from an introspective awareness of the character of experience – to metaphysics (and most recently), to a conclusion about the intrinsic nature of the brain. To Descartes, such moves would be not only unacceptable, but fallacious.

Even the self-evidently true proposition '*I am, I exist*' (AT VII.25) in the *Second Meditation*, does not license any inference to the essence of mind or to the essence of what I am, since the considerations in the *Second Meditation* are based "in an order corresponding to my own perceptions" not "in an order corresponding to the actual truth of the matter," as he explains in the *Preface to the Reader* of the *Meditations* (AT VII.8). Thus, in the *Second Meditation*, Descartes does not argue that the essence of the meditator is thinking, only that the attribute of thought cannot be separated from the meditator, because it cannot be doubted without a contradiction. In other words, his commitments there are epistemic, not metaphysical. In fact, in a very important passage invariably neglected by writers, he raises the question: "yet may it not perhaps be the case that these very things [e.g., the human body] which I am supposing to be nothing, because they are unknown to me, are in reality identical with the 'I' of which *I am aware*?" (AT VII.27; emphasis added). "I do not know, and for the moment I shall not argue the point, since I can make judgements only about things which are known to me. I know that I exist." But the formidable question is: "what is this 'I' that I know?" (AT VII.27). This exemplifies Descartes' rigorous order of reasoning: he does not argue from what he doesn't know, or from ignorance; nor does he move from epistemological considerations to metaphysical conclusions.

Descartes rejects introspection, which he calls "internal sense," because he holds it to be unreliable as a means to acquiring knowledge or grounding any metaphysical commitments (AT VII.77). His argument for dualism is thus not based on introspection, nor is it concerned with sensory experiences, nor with the sense data of the logical positivists or empiricists. Descartes' primary concern is with reason, not the senses or the imagination, and hence with knowledge based on what can be clearly and distinctly understood by reason, on what can objectively be the case. His primary and central concern is with the nature and intelligibility of *what is real and true*.

The principal attribute of a mental substance is thought. In the *Definitions* Descartes explains: "*Thought*. I use this term to include everything that is within us in such a way that we are immediately aware of it" (*Def. I*). Two points of clarification: first, by 'within us' he means within the mind, but this doesn't denote a spatial relation, only a dependence relation, since the mind is not a container; secondly, by 'immediately aware' he means that we're cognitively aware of the mind's *acts*, such as "the operations [or acts] of the will, the intellect, the imagination and the senses." By 'immediate' he means there is no gap between, say, the act of thinking and my being aware of that act — and so on for all

---

<sup>13</sup>Unlike some of the Scholastics whose defence of reason was in terms of its service to religion and faith, Descartes liberated reason from all external authorities and defended reason's authority and autonomy in its own right.

other acts of the mind. An act of thinking is distinguished from *the content* of that thought.<sup>14</sup> The content of any thought is objective and capable of being understood by any thinker, it can be common to several thinkers; there is nothing private in the Wittgensteinian sense, or ineffable, or mysterious about the content of any thought. But the act of thinking is a subjective act precisely because it pertains to a *subject*; it's performed by each subject or thinker.

The notion of objectivity has been misappropriated over the last hundred years by various thinkers, using it interchangeably with, or equivalently to physical, scientific, neuroscientific, etc. But even if everything physical is objective, the converse is not true. All our mental concepts are objective, if Frege is right about not only the publicity but the objectivity of language, even though they are not amenable to a physicalist, a scientific, a neuroscientific, or a behaviouristic analysis or reduction. The mesh of objectivity is much finer, allowing it to catch far more than the mesh of physics or of behaviourism, and its scope or extension is far wider than physical reality and behavioural dispositions.

It's true that I cannot perform your act of thinking, nor can you perform mine; but *what* I am thinking, the content of my thought, can be thought by any thinker, and *vice versa*. Willing, understanding, imagining, perceiving, sensing are faculties, not separable parts, of the mind — they presuppose the same *I* (AT VII.86; AT VII.28) and the unity of self-consciousness. Its modes are its acts, ideas, and mental states, and are indivisible. Sensations, emotions, passions, etc., also fall “under the common concept of [...] consciousness, and we call the substance in which they inhere a thinking thing or a mind” (*Third Set of Objections with Replies* AT VII.176; inherence denotes dependence not spatial relation). They're not a third category; they arise from the substantial union of mind and the human body (as we shall see later), and at least in our case as self-conscious embodied human beings or persons “there is an intellectual act included in their essential definition” (AT VII.78). Self-consciousness implies the capacity for self-attribution of those modes, the capacity for self-reflection, the ability to step back from them, a step towards adopting a more objective standpoint, and to subject them to self-governed critical evaluation and scrutiny.

The principal attribute of a corporeal substance is extension. In the *Rules for the Direction of the Mind* Descartes explains: “By ‘extension’ we mean whatever has length, breadth and depth,” leaving aside at this state “the question whether it is a real body or merely space” (*Rule Fourteen* AT X.442). In the *Synopsis*, he explains that he will develop “a distinct concept of corporeal nature,” which is “developed partly in the *Second Meditation* itself, and partly in the *Fifth and Sixth Meditations*” (AT VII.13); that is, “the whole of that corporeal nature which is the object of pure mathematics” (AT VII.71) and “of geometrical demonstrations which have no concern with whether that object exists” (French edition; CSM:49). In the earlier work, the *Discourse* Part IV (AT VI.36), he says: “I considered the object studied by geometers. I conceived of this as a continuous body.” And in the later work, the *Principles* he explains: ‘Body’, strictly, refers to “the whole universe of corporeal substance [with] no limits to its extension” (*Principles* II.21–22).

Extension and thought are simple notions, that is, indivisible both in reality and in reason. Simple notions and simple common notions (such as substance, existence, unity, duration, union, which are attributed indifferently to corporeal and thinking things) are the simplest constituents of knowledge and are known through themselves (*Principles* I.10 AT VIIIA.8). They're unanalysable, primitive, irreducible and prior in the order of reasoning; their self-evidence “is the basis for all the rational inferences we make” (*Rule Twelve* AT X.419).

It can be argued, that modern physics has superseded the corpuscularian mechanics of Descartes' time, which consisted of imperceptible particles of various shapes bumping into and hooking up with each other in various ways. Quantum field theory has done away with such particles of the past.

---

<sup>14</sup>Descartes anticipates Frege's (1948) subjective/objective distinction. Frege argues: “By a thought I understand not the subjective performance of thinking but its objective content, which is capable of being the common property of several thinkers” (p. 215 fn5).

Therefore, we have no reason to adhere to Descartes' conception of the physical which, according to recent theories of the mind, inevitably excluded the possibility of conscious experiences being physical.

It is reflection encouraged by his methodic doubt that leads Descartes to turn to a conception of a mathematically based physics. He's fully aware that any theory of physics of physical nature may be rendered corrigible by the progress of science; no doubt that's true of our modern physics, too. And it's precisely that realisation that leads Descartes to the mathematical/geometrical notion of extension, using his coordinates to express spatial positions (and thus to do, for example, geometry without appealing to the diagrams of the imagination, thereby lifting geometry out of the realm of the imagination).<sup>15</sup> Thus even if modern physics has done away with Descartes' physics, what remains part of, and indispensable to modern physics is precisely what Descartes helped to initiate and give impetus to: a mathematised physics which has remained a paradigm. As his conclusion of *The Principles of Material Things* he states: "The only principles which I accept, or require, in physics are those of geometry and pure mathematics; these principles explain all natural phenomena, and enable us to provide quite certain demonstrations regarding them" (*Principles* II.64). His aim was to show that mathematically based physics is possible and can be founded on reason, but he was not in the grip of the fantasy that reason unaided can give us knowledge of a world of corporeal objects. He considers, exactly as our current scientific enterprise does, observation and experimentation to be the spirit of scientific enquiry, and he ridicules those who ignore experience, that is, scientific observations, testing and correction (*Rule Five* AT X.380). He says: "I use no reasoning which is not mathematical and evident, and all my conclusions are confirmed by true observational data" (Letter to Plempius for Fromondus 3 October 1637 AT I.421; CSMK:63–4).

Just as Descartes' mathematically based physics is purged of the Scholastics' occult qualities, becoming a quantitative science of order and measurement, it's also increasingly inappropriate for, and inapplicable to, an investigation of a number of metaphysical questions, including the most recalcitrant of all: the mind and its nature. The primary question of his strict order of reasoning is 'what is real and true?' It's that question which directs Descartes towards, and demands an explanation of the source of the metaphysical puzzle of *what it is to be a self* and makes the question of the self and the nature of the mind unavoidable and ineliminable from his philosophical enquiry. Implicit in all this is that the mind and its nature cannot be an object of pure mathematics and of geometrical demonstrations; nor can it be represented by Cartesian coordinates.<sup>16</sup>

It's necessary to point out here (and will be relevant to our discussion of the argument for Dualism) what Descartes explicitly states in *Definition VII*: "Whether what we call mind and body are one and the same substance, or two different substances, is a question which will have to be dealt with later on" (AT VII.160–2). That is, it will require cogent arguments, and not simply definitions, even if such definitions are real and not linguistic. In fact, Descartes in his reply to Hobbes, explains: "it is perfectly reasonable, and indeed sanctioned by usage, for us to use different names for substances which we recognise as being the subjects of quite different acts or accident [properties]. And it is reasonable for us to leave until later the examination of whether these different names signify different things or one and the same thing."<sup>17</sup> We can use any term we like, he suggests, but "once we have formed two distinct

---

<sup>15</sup>Descartes' coordinates form the basis of analytic geometry and are essential tools to disciplines such as physics, astronomy, and many more.

<sup>16</sup>Our contemporary theorists of the mind who object that Descartes inevitably excluded the possibility of conscious experiences being physical, must think that the mind can be an object of pure mathematics, geometry, and Cartesian coordinates. We therefore need an argument to that effect, not simply a dismissal of Descartes' conception.

<sup>17</sup>Descartes here anticipates Frege's famous distinction between *sense* and *reference*.

concepts of these two substance [we can establish] whether they are one and the same or different” (*Third Set of Objections with Replies* AT VII.176; see also **An Evaluation of the Argument** below). In other words, Descartes makes no presumptions that mind and body are distinct, or are identical. If he had, the argument for dualism would have merely begged the question.

### A Real Distinction

What is a Real Distinction? Earlier, I considered a modal distinction that holds between substance and modes, and a conceptual distinction that holds between substance and attribute. What kind of distinction holds between substances? The distinction between two or more substances is a *real distinction*, founded in reality, and falls into two types, which can be explained as follows.

*Numerical distinctness*: a real distinction holds between numerically distinct substances. *Def. X* states: “Two or more substances are said to be *really distinct* when each of them can exist apart from the other” (AT VII.162). Numerical distinctness is necessary but not sufficient for dualism, since two numerically distinct substances can be of the same kind, or essence. Thus whether numerically distinct substances are also really distinct *kinds* of substance, is a further question and requires further arguments, as stated in *Definition VII*.

*Metaphysical distinctness*: really distinct *kinds* of substance, that is, distinct metaphysically, is what Descartes understands by “The Real Distinction between Mind and Body” — by Substance Dualism. The term ‘dualism’ is often used without consideration of the distinction between really numerically distinct substances and really distinct *kinds* of substance. ‘Dualism’ applies to the latter, to metaphysical distinctness. Metaphysical distinctness is grounded in or presupposes numerical distinctness, or *separability* and *irreducibility*. I call this the *Grounding Principle*. There are some further principles central to the argument for Dualism that need elucidation.

*The Principle of Clarity and Distinctness*: this principle was discovered and laid down in the *Third Meditation* (AT VII.35).<sup>18</sup> It’s one of, if not his most important principle. ‘Clarity’ means that a perception, an idea, a concept, a notion, a proposition “must be present and accessible to the attentive mind” (*Principles* I.45). That means that having been subjected to Descartes’ famous methodic scrutiny, it withstood that scrutiny. ‘Distinctness’ means that the understanding of a perception, an idea, a concept, a notion, a proposition is “separated from all other perceptions” (from anything extraneous to it), and “contains within itself only what is clear” (*Principles* I.45). “A concept is not any more distinct because we include less in it; its distinctness simply depends on our carefully distinguishing what we do include in it from everything else” (*Principles* I.63). Clarity and distinctness are marks of truth. What is clearly and distinctly understood is true; it cannot be contradicted and cannot be doubted — even the evil demon cannot touch it while the meditator is concentrating on it during the continuous movement of thought. Understanding clearly and distinctly is not merely thinking or considering, but directly connects with what can be objectively true and real (*Rule Twelve* AT X.421). Our clear and distinct understanding is not only constrained by the nature of things; it’s entity-directed and entity-involving, disclosing the nature of things, and thus closing the gap between true thought and reality.<sup>19</sup>

Following Descartes’ systematic order of reasoning and scrutiny, my elucidation of the *Principle of Clarity and Distinctness* demonstrates that any objections to the effect that one might be wrong in thinking that one has a clear and distinct conception miss the point and fail to grasp the rigorous demands of Descartes’ enquiry.

*The Principle of Correspondence*: in clearly and distinctly understanding, I am understanding that “truth consists in being,” the intrinsic denomination of truth (Letter to Clerselier 23 April 1649; AT V.356; CSMK:377), thus grasping a substantial fact: “whatever is true is something” (AT VII.65), not simply of something. For Descartes, what is real is true, and what is true is real. Truth is simple, that is,

<sup>18</sup>For the discovery of the principle, see Christofidou (2013, Ch. 4).

<sup>19</sup>Even in sensory perceptions, Descartes rejects the empiricists’ veil-of-perception thesis (see Christofidou, 2013).

unanalysable, indefinable, and indivisible (*Seventh Set of Objections with Replies* AT VII.548; Letter to Mersenne March 2642 AT III.544; CSMK:211). Nevertheless, it's possible to explain the meaning of the word 'truth' "to someone who does not know the language": "the word 'truth', in the strict sense, denotes the conformity of thought with its object" (Letter to Mersenne 16 October 1639 AT II.598; CSMK:139). This is the extrinsic denomination of truth, the conformity or correspondence of a true thought with reality, the true nature of things.

The *Metaphysical Conceivability Principle*: this principle provides the test for determining, for example, whether the nature of mind and the nature of body are really distinct. As we saw in 2.2, Descartes is concerned with *conceiving*, not with mere *imagining* (*Fifth Set of Replies* AT VII.364), and more strongly, he's concerned with conceiving clearly and distinctly the nature of things — the nature of things is not up to us. Clarity and distinctness pertain to metaphysical conceivability, not to any epistemic or psychological conceivability or imaginability. Nothing less than the demand of both marks of truth, clarity and distinctness, would be acceptable. Despite the rigorous demands on clarity, on its own is insufficient for basing on it, or drawing any metaphysical commitments (*Principles* I.45).

Descartes is clear about the scope of the conceivability–possibility principle and warns against its heedless use (something to which contemporary thinkers need to pay serious attention). Such a principle, he argues: "is true only so long as we are dealing with a conception which is clear and distinct, a conception which embraces the possibility of the thing in question [... and] we ought not to use this rule heedlessly, because it is easy for someone to imagine that he properly understands something when in fact he is blinded by some preconception and does not understand it at all" (*Comments on a Certain Broadsheet* AT VIII.B.352). We can see that the subjection to the methodic scrutiny is indispensable to what can be established as clear and distinct, to freeing the mind from preconceived opinions and habits. We can also see how the two principles, the *Principle of Clarity and Distinctness* and the *Metaphysical Conceivability–Possibility Principle*, draw upon one another.

### **The Argument for the Real Distinction, or Substance Dualism**

With the necessary metaphysical framework and the complex preparatory ground in place, we are ready to consider Descartes' argument for dualism, which is presented in the *Sixth Meditation* (AT VII.78). The reasoning of the argument follows "an order corresponding to the actual truth of the matter" (AT VII.8), to what belongs to substances. It doesn't follow "an order corresponding to [his] own perception," in other words, the argument is not based on epistemological, but on metaphysical considerations. Descartes starts with some general grounding premises (GP):

GP1: "I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it."

This first grounding premise draws upon the principles of clarity and distinctness, and of correspondence between thought and the true and immutable nature of things.

GP2: "The fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct, since they are capable of being separated, at least by God. The question of what kind of power is required to bring about such a separation does not affect the judgement that the two things are distinct [that is, separable]."

This second grounding premise makes two points: first, it clearly states that the mention of God is just a way of referring to logical possibility, since the argument is not about the separation of mind and body, but about their distinctness. Therefore, whatever power is required to bring about a separation

doesn't affect the argument that the two substances are distinct, numerically and metaphysically. Whether one is a theist or an atheist is irrelevant to substance dualism (see beginning of this section above).

Secondly, and relatedly, it states that the argument is about the distinctness of complete things. Two substances “are *really distinct* when each of them can exist apart from the other” (*Def. X*). If two things can exist apart, then clearly they are not one and the same thing; it's impossible for one thing to exist apart from itself. In order to show that two numerically distinct substances are also metaphysically distinct *kinds* of substance – that is, that their natures are different – requires further arguments (see *Def. VII* above).

The demand therefore is for a clear and distinct understanding that the nature of one can be understood *apart* from the nature of the other, and *vice versa*: that is, the nature of each does not involve, and does not necessitate, an understanding of the nature of the other; nor does it mean clearly and distinctly understanding the nature of one while having reasons to doubt, or being unaware of the nature of the other. Rather, what is clearly and distinctly understood is the nature of mind *apart* from the nature of body and *vice versa*; that is why it is “enough to make me certain that the two things are distinct” (GP2). Not only are mind and body numerically distinct things, they are metaphysically distinct *kinds* of thing. In order to preserve the order of reasoning, the grounding premises of metaphysical distinctness are discussed below.

With these grounding premises plus the *Principle of Clarity and Distinctness* in place, Descartes is ready to present the central argument for dualism, the *real distinction argument*:

P1: “On the one hand I have a clear and distinct idea of myself, in so far as I am simply a thinking, non-extended thing;”

P2: “And on the other hand I have a distinct idea of body, in so far as this is simply an extended, non-thinking thing.”

Conclusion: “And accordingly, it is certain that I am really distinct from my body and *can* exist without it.” (emphasis added)

First, the conclusion of the argument clearly demonstrates that a *Real Distinction* implies *separability* not separation — that mind and [my] body are really distinct, numerically and metaphysically, and *can* exist without each other. It clearly demonstrates that critics of Descartes' dualism have missed the logical distinction that he makes between *separability* and separation, and mistakenly claim that for Descartes mind and the body exist separately. The conclusion also demonstrates that neither is reducible to the other, neither constitutes the nature or essence of the other, neither is a property of the other.

The move from *body* (in P2) to *my body* (in the conclusion) draws on the explanation that Descartes offers as part of the preparatory backdrop to the argument: “It is true that I may have (or, to anticipate, that I certainly have) a body that is very closely joined to me” (AT VII.78). This anticipates the existence of the corporeal substance, the existence of which is not relevant to the argument for the real distinction, only its true nature.

The move also anticipates Descartes' defence of the substantial union between mind and the human body, arguing that, unlike corporeal substance or body, the human body (with all its appropriate functions, organs, and unity) is substantially and truly united with the mind, constituting a human being or person — a true union which is *presupposed* by their bidirectional causal interaction (AT VII.81; Letter to Regius January 1642 AT III.508; CSMK:209). I shall come back to this when I turn to examine Ryle's behaviourist objections to Descartes' dualism.

Secondly, in addition to GP1 and GP2, the argument draws on two further grounding premises (GP3 and GP4) for the metaphysical distinctness between mind and body. In the *Fifth Meditation* it was

demonstrated that the essence of corporeal substance is extension. This, consistent with Descartes' strict order of reasoning and the methodic scrutiny, is taken up and cross-examined in the *Sixth*; having withstood that scrutiny it's established that its nature or essence is extension — the mathematical/geometrical notion of extension: GP3.

By the *Sixth Meditation* the self has a clear and distinct understanding of what belongs to him: GP4 — expressed thus: “simply by knowing that I exist and seeing at the same time that absolutely nothing else belongs to my nature or essence except that I am a thinking thing, I can infer correctly that my essence consists solely in the fact that I am a thinking thing” (AT VII.78). GP3 and GP4 satisfy the demand of *Def. VII*, by providing the necessary argument for the “fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct” metaphysically, and not only numerically.

### An Evaluation of the Argument

Let's consider the argument a little more closely. The argument, drawing on GP4, demonstrates that “absolutely nothing else belongs to my nature or essence” *in so far as* I am simply a thinking thing, a non-extended thing; and, conversely, for corporeal substance, *in so far as* it is simply an extended thing, a non-thinking thing. Descartes is not defining the mind negatively (another misattribution to him by many critics). Rather, as he explains: “when I added that the mind is non-extended, I *did not* intend to explain what the mind is, but merely to point out that those who think it is extended [is material or corporeal] are in error” (AT VII.388; emphasis added). The phrase ‘*in so far as*’ is crucial here; by means of it, Descartes captures what can be understood in the *strict sense*, demonstrating that he's not making claims stronger than he intended. As he goes on to demonstrate with the substantial union of mind and the human body, I, as a human being or person, am a thinking and sentient being: *Ego sum “Res cogitans [...] & sentiens.”* (AT VII.28), thus completes the order of reasoning that began in the *Second Meditation* and endorses the unity of self-consciousness presupposed by all acts of the mind.

All this is something the self was unable to demonstrate in the *Second Meditation*, because the arguments there follow “an order corresponding to [its] own perception;”<sup>20</sup> they're epistemic, not metaphysical. As we saw earlier, Descartes doesn't argue from ignorance, or from doubt, as many commentators (starting with Arnauld; see Christofidou, 2001, but also next subsection) have claimed, and still claim. Even in the *Third Meditation* where he establishes the *Principle of Clarity and Distinctness*, and he clearly and distinctly perceives that not even the evil demon can touch the truth that he is something, he draws no conclusion regarding his *essence*, even if his enquiry thus far reveals something about his nature — that he is doubting, affirming, denying, is willing, unwilling, and also imagines and has sensory perceptions (AT VII.29).<sup>21</sup> Indeed, as late as the *Fourth Meditation*, he still raises the question: “I happen to be in doubt as to whether the thinking nature which is in me, or rather which I am, is distinct from this corporeal nature or identical with it. I am making the further supposition that my intellect has not yet come upon any persuasive reason in favour of one alternative rather than other” (AT VII.59). This is because he also needs to establish a clear and distinct conception of the nature of corporeal substance; without a clear and distinct understanding of *both* no metaphysical conclusions can be established.

---

<sup>20</sup>This is also the case with the argument in the *Discourse* (see Descartes' explanation in the *Synopsis*, AT VII.8; see also Christofidou, 2001).

<sup>21</sup>As I argue elsewhere (Christofidou, under review), this is of groundbreaking significance: it challenges the Aristotelian real distinction between the sensory soul and the intellectual mind, and it endorses a move away from that dominance and towards Descartes' conception that both intellectual and sensory acts are dependent on and are inseparable from the mind (*mens*) (AT VII.28–29), all of which presuppose “the same ‘I’” (AT VII.28).

It's in the *Sixth Meditation* that Descartes can finally draw the real distinction between mind and body, because it's in the *Sixth* that he establishes the nature of corporeal substance — GP3. The rigour of his order of reasoning is unrelenting (if not unsurpassed). It's the non-identifiability of the nature of mind with the nature of body and *vice versa*, that constrains our thinking. We then know “*the fundamental ground of difference*” (Evans, 1982, p. 107) of *that* entity as distinct from *that* other entity, and *vice versa*. Each true entity itself is “in fact the possessor of the fundamental ground of difference” (ibid.). In knowing that, it's not reason that *makes* it thus and so; rather, reason spontaneously asserts the assent-compelling truth (See Descartes' *Fourth Meditation* and Christofidou, 2013). Knowing clearly and distinctly *which* substance it is cannot leave the question open as to what *kind* it is; these are the requirements of reference and singling out based on clear and distinct understanding. It's the true nature or essence of things, their fundamental ground of difference, that constrains the true signification of the names we use (see **Kinds of distinction** above): “our mind is not the measure of reality or of truth; but certainly it should be the measure of what we assert or deny” (Letter to More 5 February 1649 AT V.274; CSMK:364). The ‘should’ indicates epistemic responsibility.

### Some Objections and Replies

There are two objections that I should like to consider here (for various other objections and Descartes' replies, see Christofidou, 2001; 2013, Ch. 10). The first is raised by Arnauld, though as is mostly the case with Arnauld's objections, he oscillates between referring to the *Second* and the *Sixth Meditations*. He argues: “[although] I can obtain some knowledge of myself without knowledge of the body [...] it is not yet transparently clear to me that this knowledge is complete and adequate [...] so as to enable me to be certain that I am not mistaken in excluding body from my essence” (*Fourth Set of Objections* AT VII.201). Arnauld's objection is rather confused to say the least, since a real distinction, as we have seen, requires knowledge of *both* entities in question, not just one.

Nevertheless, it's true that according to Descartes, we don't have an adequate conception of ourselves. Such a conception requires complete knowledge of the entity in question. Descartes' reasons are two-fold: first, as he explains to Arnauld: “There is a great difference between, on the one hand, some item of knowledge being wholly adequate, which we can never know with certainty to be the case [...] and, on the other hand, its being adequate enough [sufficient] to enable us to perceive that we have not rendered it inadequate by an abstraction of the intellect” (*Fourth Set of Replies* AT VII.221–23). We can never be certain that an idea is adequate, because that would require an enumeration of all the attributes, properties, or modes pertaining to the entity in question; such an enumeration would go on indefinitely, given that we could never know that we had reached the end. Such item of knowledge would be “an abstraction of the intellect,” rendering “it inadequate,” not clear and distinct.

Descartes continues: “when I said that a thing must be understood *completely*, I did not mean that my understanding must be adequate, but merely that I must understand the thing well enough to know that my understanding is *complete*” (AT VII.221–23). For example, when I clearly and distinctly understand what a triangle is, I cannot possibly grasp everything there is to know about it (there's an indefinite number of truths about triangles); my clear and distinct understanding is not complete (of the totality of truths) of the nature of the triangle — rather, I understand enough of the essential nature of the triangle (e.g., that it's a plane three-sided figure; if it's right-angled, I include in my understand that it has the Pythagorean theorem, etc.) for genuine understanding. Thus, it's not that I understand the complete nature of the triangle, but that I completely understand the nature of the triangle, and also that all other truths about the triangle cannot be inconsistent with its nature.

Secondly, and relatedly, in his letter to Gibieuf (January 19, 1642), Descartes says: “I do not deny that there can be in the soul or the body many properties of which I have no ideas; I deny only that there are any which are inconsistent with the [clear and distinct] ideas of them that I do have, including the ideas that I have of their distinctness” (AT III.477–78; CSMK: 203). However many other properties might belong to the mind or to the body of which he has no knowledge, he's not committed to knowing *a priori*, or *a posteriori* all properties or modes of each substance. What is essential is that they cannot be inconsistent with the principal attribute “which constitutes [the substance] nature and essence, and



to which all its other properties are referred" (*Principles* I.53). If they are inconsistent, they are excluded; this is because a true entity cannot have inconsistent properties entailed by its true nature.

The second objection, which has persisted over the centuries, and which has been deemed to be insoluble, is known as the mind–body problem. The mind–body problem primarily involves the question of causality; that is, the causal interaction between the mind and the human body. As I have recently argued for a solution to the mind–body problem, I shall restrict my discussion here to two points:

First, central to my defence is an account of Descartes' metaphysics of causality. Causality is a metaphysically basic category, along with such basic common notions as substance, essence, unity, truth; basic common notions, as we have seen, are attributed indifferently to mental and corporeal entities. Causality can be given an elucidation (as Descartes does in the *Third Meditation* where he presents his *Causal Principle*) but it is unanalysable and irreducible (for a detailed elucidation and discussion of Descartes' *Causal Principle*, see Christofidou, 2019). Notions such as force, contact, impact, transfer of energy, interaction, or power *presuppose* causality; they're causal notions, and thus cannot "provide a basis for an account of causality," contrary to what some philosophers argue (Gibb, 2015, p. 140), nor can "causality just [be] a manifestation of these powers" (Martin 2008, secs. 5.3, 7.4). Causality is itself neither mechanical nor non-mechanical, neither physical nor mental; the notions of mechanical, non-mechanical, physical, or mental interactions make appeal to causality, but causality itself remains prior and unanalysable. Causality itself, and Descartes' *Causal Principle*, make no assumptions, *a priori* or empirical, about the nature of the causal relata. The neutrality of the *Causal Principle* bridges the metaphysical distinctness of the two substances.

Secondly, Gassendi and other commentators, past and present, assume, without independent arguments, that causation is a physical relation, and hence that heterogeneous interaction puts pressure on the intelligibility of mind–body interaction, or is inconsistent with Descartes' *Causal Principle*. Descartes, in his replies to Princess Elizabeth (in his various letters) and to Gassendi's *Counter-Objections* regarding the mind–body relation, states: "the whole problem contained in such questions arises simply from a supposition that is false and cannot in any way be proved, namely that, if the soul and the body are two substances whose nature is different, this prevents them from being able to act on each other" (AT IXA.213).

Descartes didn't think there was a mind–body problem, because he didn't make the unfounded physical causal closure assumption, nor did he assume that causal relata must be homogeneous, or that reality is physical — that the limits of one's ideology are the limits of reality. Rather, his struggle was to explain to his critics his new metaphysics (which was challenging their Scholastic conceptions) by offering a way of understanding both the mind–body relation and the uniqueness and presuppositionality of the mind–body union, and by pulling his critics back from their misconceptions about the metaphysics of causality and of his *Causal Principle*. For Descartes, as it was for the classical philosophers, Plato and Aristotle, and for Spinoza, Leibniz, and Russell, among others, philosophy is concerned with the universe as a whole, and the universe is not simply the physical, neurobiological, or dispositional behavioural world.

### **Philosophical and Logical Behaviourism**

In order to put Gilbert Ryle's attacks on Descartes' dualism into context, a brief characterisation of behaviourism is in order. Behaviourism came into being at the turn of the twentieth century as a methodological research programme in experimental/scientific psychology. The supposition was that something about the mind prevented psychology from attaining the authority and status of the physical sciences, such as chemistry or biology. The concern, therefore, was that, if psychology were to be scientific, it had to deal with publicly observable and measurable behavioural phenomena, operating in

the space of scientific causes and the realm of laws. This methodological behaviourism set aside ontological or metaphysical questions – that is, questions about the existence or nature of the mind, consciousness, etc. – in order to proceed with what it took to be a scientific approach: behavioural tests, prediction, and control. It made no proposals about how consciousness is or might be related to behaviour. As Solomon (1993) explains, J. B. Watson’s “own crude and extreme restrictions on the appropriate data and conceptions of psychology [...] limited the range of intelligible and interesting statements that psychologists could make about *human* behavior.” From this “it is a small step to the suggestion that consciousness can be *ignored* in the study of behavior, and an even smaller step, though replete with metaphysical atrocities, to the thesis that there is no need to talk about ‘consciousness’ in psychology at all.” It doesn’t follow that all psychology was “behaviouristic; but it must be admitted that even those who staunchly opposed behaviourism were intimidated by the well-established argument against the primacy of consciousness” (pp. 77, 84–85). Methodological behaviourism appeared to be non-committal with regard to philosophical implications.

When philosophical behaviourism came on the scene, it didn’t set aside ontological or metaphysical issues for methodological or scientific reasons, because ‘setting aside’ suggests that one might still adhere to what Ryle (1949) calls “the ill effects of the myth” of dualism, “the two-worlds story” (p. 310). Philosophical behaviourism, associated with Ryle, is itself a metaphysical thesis (albeit an impoverished one), according to which minds, consciousness, and all mental states are not “identified with anything, either physical or nonphysical,” nor are they even constituted by behavioural dispositions, which might suggest a constitutive thesis of the nature of mental states. Rather, mental states *are* the (covert or overt) behavioural dispositions themselves, and nothing more than publicly accessible and observable occurrences, thus deracinating the ill effects of the myth, the two-worlds story.

In turn, through conceptual or linguistic analysis, mental concepts were said not to be referring expressions — not to refer to anything mental; as Ryle states: “the cardinal mental concepts have been credited by philosophers [...] with the wrong sorts of logical behaviour [...] these concepts have been misallocated” (310). The claim is that all locutions that contain mental-concept terms, such as Ryle’s ‘the concept of mind,’ can be analysed without remainder into statements containing only terms referring to behaviour and behavioural dispositions; for example, ‘the concept of mind’ does not refer to any so-called occult or ghostly stuff, but is analysable into a description of behavioural dispositions.

Philosophical and logical behaviourism adopts an anti-mental outlook and aims to show that the mind *is* the behavioural dispositions, or forms of behavioural organisation, and that mental concepts are analysable and definable in terms of behavioural utterances. It is a type of eliminative thesis, ontologically and conceptually.

### **Ryle’s polemic against Descartes’ Dualism**

The term ‘polemic’ is more apt than ‘critique’ for reasons that I shall explain below.<sup>22</sup> Ryle coined three catchy slogans: ‘Descartes’ Myth,’ the ‘Category Mistake,’ and the ‘Dogma of the Ghost in the Machine.’ These slogans have had a lasting effect and have become entrenched in philosophical habits. I shall consider these slogans in turn, though they are bound up together.

My discussion will take the form of ‘*Objections and Replies*,’ prompted by two considerations. First, the need to answer the question: ‘who is distorting our ordinary language, violating the logical geography of concepts, committing a category mistake, and systematically misleading us?’ Secondly, the presentation of the two thinkers in dialogue will enable us to consider some serious concerns about

---

<sup>22</sup>Ryle (1949, pp. 10–11) acknowledges that some readers may think of his approach in his book as “excessively polemical,” but tries to justify this by saying that “the assumptions against which I exhibit most heat are assumptions of which I myself have been a victim.” As we shall see, such assumptions are ones that Ryle himself had conjured up. My adoption of the term ‘polemic’ is, therefore, not an excuse for exhibiting heat, but to point up the lack of engagement with any arguments of the thesis under attack.

personhood and the self, about the metaphysics of mind, the metaphysics of freedom, and moral significance. It will enable us to appreciate the immense difference between Descartes and Ryle in their approach to proper philosophical enquiry and analysis.

Ryle's declared purpose is "to prove that [Descartes dualism] is entirely false, and false not in detail but in principle" — it commits "a category-mistake" (p. 17), and hence on its own terms is logically untenable. Descartes' work abounds in subtle but important distinctions, in arguments, in a variety of complex notions, and in a number of principles. But in place of a philosophical engagement with any of these, or with any of Descartes' work, Ryle explicitly states that he prefers to "speak [...] with deliberative abusiveness" (p. 17).

My aim is to show, not only that Ryle's slogans are misplaced, but that any claim in his work against Descartes' dualism is false. I say 'any claim' because Ryle offers no arguments against *Descartes'* position. This also explains my choice of the term 'polemic' instead of 'critique,' because a critique requires at least some engagement with, and an attempt at understanding, one's opponent's arguments. Ryle, however, engages with none of Descartes' arguments, and provides no references to his work. Instead, it's through abusiveness that he aims to show not simply that dualism is false in detail, but that the very idea of it is false. Instead of trying to understand what Descartes is saying, and how his arguments are meant to work, Ryle is happy to make untutored claims about the object of his derision, usually with the aim (or, at least, with the effect) of setting up a straw man to be knocked over. The idea that one of philosophy's undertakings is the search for truth seems to have no place in Ryle's 'logical analysis' which, as he openly declares, has a "destructive purpose" (p. 19).

### Descartes' Myth

Ryle states: "Descartes left as one of his main philosophical legacies a myth which continues to distort the continental geography of the subject" (p. 10), labelled "Descartes' Myth" (p. 13).

Ryle begins his attack by stating: "There is a doctrine about the nature and place of minds which is so prevalent among theorists and even among laymen that it deserves to be described as the official theory" (p. 13). Although the 'official theory' has, according to Ryle, a long history spanning Plato, Aristotle, Augustine, the Stoics, the Scholastics, to Descartes and beyond, the main culprit and the focus of attack is Descartes who, according to Ryle, "was reformulating already prevalent theological doctrines of the [immortality of the] soul" (p. 24).

What seems to have escaped Ryle is the fact that Descartes makes it explicit that he was not concerned with theological doctrines, let alone reformulating them. Indeed, as we saw earlier, one of his aims, and eventual achievements, was the divorce of theology from philosophy, which he saw necessary in order to enable the emerging new sciences to progress. Descartes liberated reason's power from external authority and enabled the will to free itself from the shackles of prevailing entrenched doctrines, to assert its autonomy and be at one with reason, not pulling in opposite directions (see Christofidou, 2013). Reason's authority and freedom's autonomy in their internal relation was a necessary condition of his search for the possibility of a new metaphysics.

Moreover, Descartes states in a number of places that he doesn't (and couldn't) argue for the immortality of the soul, including the *Synopsis* to the *Meditations* (AT VII.13–14; see beginning of the second section above). It can also be seen from the objections, raised by a number of theologians,<sup>23</sup> that

---

<sup>23</sup>The *Second Set of Objections* is "attributed to 'theologians and philosophers' in the index to the first edition [1640], but the French version of 1674, announces that they were 'collected by the Reverend Father Mersenne'. In fact, they are largely the work of Mersenne himself" (*Translator's Preface* vol. II:64).

Descartes said, “not one word about the immortality of the human mind” (*Second Set of Objections* AT VII.127), to which Descartes replied: “from the fact that the soul is distinct from the body [it doesn’t follow] that it is immortal.” It’s possible, he said, “that its duration comes to an end simultaneously with the end of the body’s life.” (*Second Set of Replies* AT VII.153. See also the *Fourth Set of Replies* AT VII.248; *Sixth Set of Replies* AT VII.428).

The theologians, who had read Descartes’ works, clearly saw that he didn’t say a word about the immortality of the soul,<sup>24</sup> while Ryle, who seems not to have read Descartes’ works, simply asserts that Descartes was reformulating theological doctrines of the immortality of the soul. This introduces a pattern which is repeated more than once.

Ryle continues: to determine “the logical geography of concepts is to reveal the logic of the propositions in which they are wielded, that is to say, to show with what other propositions they are consistent and inconsistent, what propositions follow from them and from what propositions they follow” (p. 10). Descartes’ propositions, he claims, distort this “logical geography of concepts” — such as the proposition: “every human being is both a body and a mind” (p. 13). Ryle states: “It is perfectly proper to say, in one logical tone of voice, that there exist minds and to say, in another logical tone of voice, that there exist bodies. But these expressions do not indicate two different species of existence, for ‘existence’ is not a generic word like ‘coloured’ or ‘sexed’. They indicate two different senses of ‘exist’” (p. 24). Thus, Ryle claims, Descartes is committed to two different species of existence in the proposition ‘every human being is both a body and a mind’, which distorts the “logical geography of concepts”.

How cogent are these claims? First, for Descartes, the notion of “existence [along with the notions of] unity, truth, duration and the like” is a “simple [basic] common notion”, ascribed indifferently to bodies, minds, and so on (*Rule Twelve* AT X.419). The concept of existence is univocal. Ryle offers no argument against this, and no argument for his own position concerning existence. In fact, secondly, and more generally, it’s not clear that ‘exists’ can be used to indicate either two species or two senses of existence. Rather, to say, ‘Hamlet exists’ and ‘Ryle exists’ is to refer to two different domains of discourse; ‘exist’ has to mean the same in each case, just as ‘... lived in Denmark’ has to mean the same in ‘Hamlet lived in Denmark’ as in ‘Niels Bohr lived in Denmark’ — otherwise we wouldn’t be able to understand fiction. In the fictional domain of discourse, Hamlet exists and thinks. In the domain of discourse of the real world, Ryle exists and behaves. For Descartes, given that the mind and the human body are both true natural entities, and *are* substantially united, a union “ordained by nature” (*Optics* AT VI.130), they exist in the same domain of discourse, the real actual domain. Ryle is not meeting Descartes on his own terms; instead, he’s simply presupposing his behaviourist thesis, which denies the existence of minds as true natural entities, and denies that mental concepts refer to anything mental.

Ryle illustrates the notion of a distortion of the “logical geography of concepts” with a number of examples of zeugma, such as: “‘She came home in a flood of tears and a sedan-chair’ [which is an] absurdity of conjoining terms of different types” (p. 23). He then claims that what he calls Descartes’ Myth, “the dogma of the Ghost in the Machine does just this. It maintains that there exist bodies and minds” (p. 23) the conjunction of which violates “the logic of concepts” and the logic of the sentences in which the concepts are used. Hence dualism is on its own terms logically untenable. Even if, Ryle says, he is not “denying that there occur mental processes,” nevertheless, “the phrase ‘there occur

---

<sup>24</sup>In some of his letters to friends, such as to Constantijn Huygens, whose brother had died, Descartes offers some consolatory thoughts: “Those who die pass to a sweeter and more tranquil life than ours; I cannot imagine otherwise. We shall go to find them some day, and we shall still remember the past; for we have, in my view, an intellectual memory which is certainly independent of the body” (10 October 1642 AT III.798; CSMK:216). Similarly, in his letter to Princess Elizabeth, who was going through a time of unhappiness, Descartes wrote: “even without the teachings of faith, natural philosophy by itself makes us hope that our soul will be in happier state after death than now” (1 September 1645 AT IV.282; CSMK:263). But as Clarke (2003) argues, these were prevalent theological doctrines about “the kind of happiness” in afterlife “rather than a conclusion derived from the Cartesian theory of mind” (p. 101), or from Descartes’ metaphysics.

mental processes' does not mean the same sort of thing as 'there occur physical processes,' and, therefore, [...] it makes no sense to conjoin or disjoin the two" (p. 23). The conjunction of these phrases "and other analogous conjunctions are absurd" — for the reasons just given.

The central problem here is that Ryle hasn't thought carefully enough about what kind of absurdity is involved, and where in a zeugma it lies. First, zeugmas such as: 'She came home in a flood of tears and a sedan-chair' are not absurd in the logical sense; on the contrary, they work precisely because we understand exactly what they mean. The mismatch between grammar and logical form doesn't give rise to logical absurdity but to comical absurdity. Secondly, the humorous effect of a zeugma comes not from the mere conjoining of two different sorts of thing (as Ryle maintains), but in the *manner* of their conjunction: the word doing the work in the example is 'in,' which means something very different when attached to 'a sedan chair' (spatial position, containment) and to 'floods of tears;' the absurd effect is achieved by using a single token of 'in' to govern the two noun phrases. Thus, if someone were to say: 'she sat in a sedan chair in floods of tears' (or even 'she returned home in a sedan chair and in floods of tears'), we would find little that was humorously absurd. There is nothing like this either in Descartes' substance dualism, or in his substantial union of a mind and a human body.

Thus Ryle's claim that Dualism is logically absurd after the manner of a zeugma fails on two counts: first, he has misunderstood the way in which zeugma works, nor did he notice that it doesn't apply to the target of his attack; secondly, even if Descartes' position had been an example of zeugma, this would only have suggested some humorous absurdity, not the logical absurdity that Ryle claims to have demonstrated. His discussion here is again polemical rather than rational; in a familiar move, Ryle distracts us with the amusement of examples of zeugma, apparently in the hope (unfortunately often fulfilled in the case of many readers<sup>25</sup>) that we'll not notice that his slide into attacking Dualism is unargued. We also see another repeated tactic: Ryle is prone to prefixing and peppering his claims with terms such as 'logical absurdity,' 'logically fatal,' 'logically trim,' 'logical mould,' 'logical muddles,' and with phrases such as 'in one logical tone of voice,' which either don't mean anything or are misused, but whose purpose is to serve as sticks to browbeat the reader.

Indeed, let us consider one of Ryle's own usages; take 'there is a ghost,' which clearly does not mean the same sort of thing as 'there is a machine,' in that the former involves a fictitious entity while the latter refers to a physical object. We must conclude that Ryle's conjunction 'there is a ghost and there is a machine, and the ghost is in the machine,' at the core of his polemic, is by his own lights absurd, violating the logic of the concepts, and the logic of the sentence in which the concepts are used. Equally importantly, by Ryle's own lights, the only logically legitimate sentences must be ones whose conjuncts (or disjuncts, and so on) mean the same sort of thing and can be substituted for one another without rendering the sentence absurd. That means only identity statements are logically legitimate; only if '*a* is identical to *b*' can '*a*' mean the same sort of thing as '*b*,' and be truly substituted for '*b*' in any true statement about *b*, preserving the truth-value of the statement. Any other statements are logical violations. But this is again nonsense, just as it is nonsense to claim, as Ryle does, that the conjunction "there exist bodies and minds" is absurd.

### Category Mistake

---

<sup>25</sup>This class is broad, and includes readers from different forms of behaviourism (such as Skinnerians or Kantorians), who uncritically cite Ryle's *The Concept of Mind* in support of their theses and their rejection of Descartes' dualism.

At the centre of Ryle's polemic is the 'Official Doctrine', that is, 'Descartes' Myth', of which he says he will "often speak with deliberate abusiveness as 'the dogma of the Ghost in the Machine'" (p. 17). Ryle's aim is "to prove that [Descartes' dualism] is entirely false, and false not in detail but in principle. It is not merely an assemblage of particular mistakes; it is one big mistake, and a mistake of a special kind: a category-mistake. It represents the facts of mental life as if they belonged to one logical type or category [...], when they actually belong to another. The dogma is therefore a philosopher's myth" (p. 17).

Ryle does not define what a category is, nor does he offer a theory of categories that would enable us to avoid mistakes; it's, therefore, not clear to which category the facts of mental life belong and to which they don't belong<sup>26</sup>. The question of categories (or principal distinctions of what there is, or classes, types, or genera) is a primary concern of metaphysics, and involves several interconnected questions: what are categories? What categories are there? Does a theory or a table of categories need elucidation, if not justification? Are categories immutable? It is precisely on these grounds that Warnock (1958) objects to Ryle's use of the phrase 'category mistake,' asking: "if one is not prepared and indeed is deliberately unwilling, to say just what a category is, and what categories there are, can one really be entitled to employ the term 'category'?" (p. 96).

Ryle, unwilling to define what a category is, relies on the notion of absurdity in order to determine category differences, preferring to "indicate what is meant by the phrase 'Category Mistake'" (p. 17) using not one example but several — perhaps in an attempt to convince himself, as if he were "to buy several copies of the morning paper to assure himself that what it said was true" (Wittgenstein, 1953, §265). The best-known of his examples involves a foreign visitor to Oxford, who is shown round all the colleges, libraries, museums, laboratories, and so on, and then asks: 'But where is the University?' The visitor has, claims Ryle, committed a category mistake; he assumed that the term 'University' "stood for an extra member of the class of which these other units are members. He was mistakenly allocating the University to the same category as that to which the other institutions belong." But of course, once all the institutions "are seen and their coordination is understood, the University has been seen" (p. 18). There is nothing over and above the members of the class. Similarly, Ryle states, dualism assumes that the term 'mind' stands for an extra member of the class of publicly observable behaviour and behavioural dispositions; but once the dispositions and behaviour are seen and their coordination is understood, the mind has been seen. There is nothing other than the behaviour and behavioural dispositions.

We find here another evidence of Ryle's approach that we discerned above, sliding from one notion of absurdity to another; he distracts and mesmerises the reader by waving around his amusing examples in an attempt to disguise the fact that he offers no argument for lumping the target of his criticism in with the examples. The fact that the examples themselves are not all accurate<sup>27</sup> is essentially

---

<sup>26</sup>Place (1999 — himself a mind-body identity theorist regarding certain mental processes, such as sensations, and a behaviourist regarding propositional attitudes, such as beliefs) thought that Ryle's work had not been more influential because of Ryle's "failure to sharpen up the notion of category mistake." (p. 371). If only that had been Ryle's sole failure!

<sup>27</sup>The University example, for instance, is flawed. As was very well known to Ryle, the University of Oxford is more than, and hence not reducible to, the things that the visitor has been shown. For example, when one is admitted to study at Oxford one is admitted both by a College and by the University; it is possible as a graduate student to gain a College place but not a University place. The University has staff who are often not staff of any of the Colleges, libraries, laboratories, etc. Colleges and University are, administratively and legally, separate, distinct organisations; however unlikely, a College might choose to leave the University, and the University might rule that a College be no longer entitled to enter its undergraduates for University examinations. Thus, when one speaks of the University one speaks of something more than the various colleges, libraries, and so on, and their coordination, just as when speaks of the mind one speaks of something more than bodily behaviour and behavioural dispositions. The mind is real and so is bodily behaviour; the University is real and so are the Colleges. It's therefore Ryle who commits a category mistake, not the foreign visitor.

irrelevant; the central point is that Ryle offers no argument for holding that the mind and body are analogous to (his misrepresentation of) the university and the colleges, etc.

In addition, there's a serious misunderstanding that affects this area of enquiry: the notion of human behaviour presupposes mental states such intentions, motives, desires, fears, thinking, judging, unlike the movements of planets or machines. If the mind is irrelevant to, or severed from human behaviour, what exactly are we considering or discussing? We seem to be considering the body of a human being, as Descartes says, as no more than "a kind of machine equipped with and made up of bones, nerves, muscles, veins, blood and skin in such a way that, even if there were no mind in it, it would still perform the same *movements* as it now does in those cases where *movement is not under the control of the will or, consequently, of the mind* [but occurs merely as a result of the disposition of the organs – French version]" (*Sixth Meditation AT VII.84*; emphasis added).

Ryle then turns to considerations of ordinary language. He claims that Descartes' dualism commits a category mistake by misusing our ordinary language; by misunderstanding how ordinary language functions, Descartes is led to ontological mistakes by postulating entities, such as the mind, that in fact do not exist. There is a logical absurdity in claiming that the terms 'mind,' 'mental, or 'consciousness' refer to something inner and unwitnessable, "since their supposed objects are myths" (Ryle, 1949, p. 149). For example, the concept pain is analysable in terms of behavioural utterance and behavioural dispositions. When we see a person holding his or her chin, we suppose that there is something going on inside the head of the person, and we use the concept pain to refer to his or her mental state of pain. But the concept pain and the statement 'he or she is in pain,' or 'I am in pain,' mean nothing more than 'such and such behavioural dispositions are going on.' And this is the case for every other mental concept and every other statement containing mental terms or concepts. There is no further mental state called 'being in pain.' To argue otherwise is to misuse our ordinary language, to violate the logic of concepts.

According to Ryle, the mind cannot be the cause of the behavioural dispositions because the mind *is* the behavioural dispositions displayed. On Ryle's philosophical behaviourism, the idea of mental causation is a category mistake, and a logical distortion of ordinary language. On this view there can be (not only in practice, but in principle) no such thing as unrequited love, suffering in silence, being in pain, being happy, or thinking without behaviour or behavioural dispositions. The outward expressions of happiness, or pain, are not due to some mental inner maneuverings inflicted on our bodies. The outward expressions themselves *are* happiness or painfulness. Conversely, there's no way of my exhibiting such behavioural dispositions but in fact being very unhappy, or not being in pain. There is nothing *it is like to be* in a particular mental state, nothing *it is like to be* a conscious being, nothing *it is like for* the conscious being since consciousness is simply mythical. Do I need to observe my behaviour in order to know that I am suffering, or thinking? But if there is no behaviour, I cannot be suffering, or thinking. And this is the case for every other mental state; the mental just is the behavioural dispositions displayed. And this is the case for every other statement containing mental concepts. To argue otherwise is to misuse our ordinary language.

But who is misusing ordinary language, Ryle, or Descartes? One asks (with Ross, 1930, p. 82): if reason has made such egregious errors about *itself*, what is it that reassures Ryle (or any other behaviourist, eliminativist, physicalist, neuronist, and so on) that reason can be trusted in its power to know anything at all, or to convince us that Ryle's polemic is true and dualism is false? Raising the worry from a different angle, Broad (1925) says: "I would invite the Behaviourist to explain *how*, on his own theory, we can ever have come to make the mistake which he says that we do make. If in fact we can observe nothing but bodily behaviour in ourselves and others, how did we ever come to entertain the hypothesis that there is something more than these; and how did we come to suppose that there are better *grounds* for assuming the presence of this extra factor in some cases than in others?" (p. 615).

Indeed, what does it mean to say that I know anything, or that I observe my behaviour, or that I believe any of Ryle's claims, given that knowing, observing, and believing *are* just dispositions to behave? Are such dispositions to behave about dispositions to behave, and so on *ad infinitum*? How can dispositions be *about* anything?

As Descartes says: "I have never seen or perceived that human bodies think; all I have seen is that there are human beings, who possess both thought [mind] and a body" (*Sixth Set of Replies* AT VII 444). But on Ryle's view, not even human bodies can think, because there is no such thing as thinking over and above the behavioural dispositions.

Descartes preempts all this when he argues: "To say that thoughts are merely movements of the body is as perspicuous as saying that fire is ice, or that white is black; for no two ideas [concepts] we have are more different than those of black and white, or those of [bodily] movement and thought" (Letter to Mersenne 21 January 1641 AT III.285; CSMK:169). And again: "if anyone asserted that Bucephalus was Music, there would be every point in someone else saying that this was false" (*Fifth Set of Replies* AT VII.388). "It is a conceptual contradiction to suppose that two things which we clearly perceive as different should become one and the same (that is intrinsically one and the same, as opposed to by combination); this is no less a contradiction than to suppose that two things which are in no way distinct should be separated" (*Sixth Set of Replies* AT VII.444) — for example, Hesperus and Phosphorus.

It's a contradiction to assert, as Ryle does, that the mind *is* the behavioural dispositions. It's as if, *mutatis mutandis*, Ryle were to assert that Alexander the Great's horse Bucephalus *is* music — a misuse of ordinary language, a distortion of the "logical geography of concepts," and a category mistake committed by Ryle. Descartes, unlike Ryle, through his cross-examination and scrutiny, provides arguments for his clear and distinct understanding of mind and his clear and distinct understanding of body. Descartes' use of language and concepts requires no rectification of their 'logical geography' from Ryle, and no category mistake is committed.

Ryle deracinates something essential: the mind and mental states and phenomena that are undeniably evident in any self-conscious subject — in any human being or person, in any human action and behaviour. His view cannot even be labelled third-personal, because the first-person and the third-person points of view are reciprocal (despite philosophers, so often, equating the third-person point of view with a scientific or a physicalist point of view).

Ryle's "destructive purpose is to show that a family of radical category-mistakes is the source of the double-life theory [Descartes' dualism]. The representation of a person as a ghost mysteriously ensconced in a machine derives from this argument" (p. 19).

### **The Dogma of the Ghost in the Machine**

According to Ryle, one's body, like every other physical object in the world, is publicly observable, while dualism claims that one's mind "is not witnessable by other observers; its career is private," from which Ryle concludes: a "person therefore lives through two collateral histories, one consisting of what happens in and to his body, the other consisting of what happens in and to his mind. The first is public and the second private" (p. 13). This is "Descartes' Myth", the "dogma of the Ghost in the Machine." It is comparable to someone who "continues to think of the Average Taxpayer as a fellow-citizen, [and thus] will tend to think of him as an elusive insubstantial man, a ghost who is everywhere yet nowhere" (p. 19).

I shall come to the single most damning response to this most carelessly polemical of Ryle's claims below. First, though, it's worth pointing out that, as usual, he shows himself to be good at devising slogans and catchy labels, while providing no references to Descartes' work to show from where any such views might be discerned. I don't know who coined the terms 'inner' and 'outer' with their specific connotations of private and public, but such terms are not part of Descartes' conception, nor are they relevant either to his dualism or to the substantial union of mind and the human body. On Descartes'



account, to argue that the mind is *inside* or *inhabiting* the body, or that the mind is inner and the body is outer would be nonsensical, since the mind and the human body are substantially united.

Ryle claims that the “phrase ‘in the mind’ can and should always be dispensed with. Its use habituates its employers to the view that minds are queer ‘places’, the occupants of which are special-status phantasms” (p. 40). It’s not clear why anyone would think that minds are queer ‘places’, but contrary to Ryle’s claim, Descartes explains that, for example, “the phrase [...] ‘it is conceived in the mind’ [...] means simply ‘it is thought of’” (*Seven Set of Objections with Replies* AT VII 480); it denotes no place or a container. When Descartes argues, for example, that attributes inhere in substances, he uses terms such as ‘inhere’ or ‘inherence’ and prepositions such as ‘in’ with their original Latin meanings, often denoting ontological, metaphysical, and explanatory dependence/independence relation (see Christofidou, 2013, for a detailed discussion), and not always a spatial relation as Ryle mistakenly assumes.

Ryle repeatedly claims that, according to the ‘dogma of the ghost in the machine,’ mental states are “occult or ghostly” (p. 33), inaudible, invisible, neither spatial, nor observable; one’s mind “is not witnessable by other observers; its career is private.” We’re therefore tempted to say that for an audible or visible behaviour there must be “some extra secret acts.” But the visible behaviour, according to Ryle, not only is “not an occult or ghostly happening, but [...] is not a happening at all. It is a disposition, or complex of dispositions” (p. 33). But aren’t dispositions inaudible, invisible, non-spatial, not witnessable by others? In which case, must they not also be accounted occult or ghostly? Ryle might respond that dispositions are witnessable by others in their manifestations in behaviour. But isn’t that equally true of the mind? One’s mind is witnessable by others in one’s purposive and intentional acts and actions, in one’s meaningful judgements, decisions, in one’s taking responsibility; responsibility is not always assignable by others. The mind is far better comprehended and understood by others than any bundle of behavioural dispositions, whose understanding and differentiation requires a *presupposition* of our mental concepts, despite being derided by Ryle. All our mental acts and states can be ascribed both from the first-person and the third-person perspective and are capable of being understood and referred to using our objective mental concepts, without insisting that they can be understood physicalistically or behaviouristically, when in fact they cannot.

Ryle’s answer is that “a disposition is a factor of the wrong logical type to be seen or unseen, recorded or unrecorded” (p. 33). He offers no argument for his claim. The question is: why couldn’t the same be said of mental states — they aren’t of the logical type to be seen or unseen? Ryle doesn’t consider this; instead, he reiterates that the ‘official doctrine’ misconstrues “the type-distinction between dispositions and exercise into its mythical bifurcation of unwitnessable mental causes and their witnessable physical effects” (p. 34).

Ryle continues: “The distinction between [...] thinking what one is saying and merely saying, cuts across the distinction between talking aloud and talking to oneself” (p. 35). Ryle treats both as equivalent, involving nothing other than the behaviour of talking; even soliloquy is nothing more than talking to oneself, and hence nothing more than behavioural dispositions. Doesn’t talking meaningfully either to oneself or to others, presuppose a mind? Is talking to oneself a model for thinking, for thinking imaginatively and creatively, for reflecting, cross-examining, doubting, and so on? And how are we to distinguish between these activities, given that none of them is anything but dispositions to behave? How are we to distinguish between dispositions of deliberately breaking a vase, and dispositions of accidentally breaking a vase? Given that any putative mental concepts are reducible to the logic of behaviour and are behaviourally definable, and given that there are *only* behavioural dispositions, how are we distinguish between thinking and reasoning, between conceiving and imagining, between reflecting and doubting or cross-examining, between discerning and discursive reasoning? How are we to distinguish between a judgement that one proposition is preferable to another and a judgement that

that proposition is true, or a judgement that it would be good if that proposition were true? What, if any, are the discriminating factors among all those behavioural dispositions? How do we identify which are which without appealing to something more than behavioural dispositions — to mental states and mental concepts that refer to those mental states, thus bringing them in through the back door, since without them nothing would make any sense?

### Ryle's Own Category Mistakes of Numbing Grossness<sup>28</sup>

I have already demonstrated that it is Ryle who commits category mistakes. The aim here is to take a closer look at and to assess Ryle's twofold accusation: that of the category mistake and that of the dogma of the ghost in the machine.

#### Ryle's own category mistakes

In the opening sentence of his *The Concept of Mind*, Ryle states that he offers “what may with reservations be described as a theory of the mind” (p. 9). One suggestion is that the mind might be thought of as a variety of “black box,” though insists “upon an absolute void of content:” no hypotheses of mental states, no mind, no consciousness, no inferences to unseen feelings, sensations, or thoughts, “just independent and dependent variables with nothing in between” (Solomon, 1993, p. 84, fn\*). But in fact, it would be misleading to use the term ‘box,’ because for Ryle there isn't even a box, which analogically would be the mind, albeit utterly empty. His behaviourist thesis cannot even be elevated to “the *hydraulic model*” which views the human mind “as a cauldron of pressures demanding their release” (Solomon, 1993, pp. 78 & 81, respectively) in behaviour.

It's not that Ryle doesn't bother with the mind (as someone might suggest in his defence) because his concern is to deflate the idea that there are inner mental processes over and above public behaviour. Nor is it that he's not concerned with analysing the inner, but only to deride it. Ryle thinks (or, should I say, is disposed to behave?) that his behaviourism doesn't reduce what is observable only to sequences of muscular impulses or motor reactions (perhaps that's why he denied that he was a behaviourist), but it reduces all mental states (and mental concepts) to dispositions to behave without remainder. It might also be suggested in his defence, that Ryle was not interested in *explaining* behaviour, only in describing it. It is, however, one thing to be concerned simply with describing behaviour; it is quite another to claim that all there is *is* behaviour and behavioural dispositions, as Ryle insists. Ryle's thesis is an ontologically reductionist purified version of the hydraulic model; in fact, it's neither a model nor a theory of the mind, because the mind and mental states are not even constituted by behavioural dispositions.

Ryle repeatedly states that there are only the behavioural dispositions, saying that it's not merely that the mind is “not an occult or ghostly happening, but that *it is not a happening at all*. It is a disposition, or complex of dispositions” (p. 33; emphasis added), and: “Overt intelligent performances are not clues to the workings of minds, *they are those workings*” (p. 57; emphasis added). Again and again, he states that it's logically absurd to claim that the terms ‘mind,’ ‘mental,’ ‘consciousness,’ and so on, refer to something inner and unwitnessable, “since their supposed objects are myths” (p. 149). Does this amount, even “with reservations,” to a philosophical or even psychological “theory of the mind?” Does it even amount to an account of *human* action, given the impoverished conception of human behaviour?

One of many things that seems to have been overlooked by Ryle is that dispositions, habits, skills, tastes, inclinations are contingent, because they can be lost, abandoned, got rid of, or acquired — we may even fail to exercise inherent capacities. But consciousness stands, not in a contingent relation, but in a special and essential relation to the subject of experience, behaviour, and thought, a subject that is a necessary presupposition for any conscious states — despite Ryle's derisive comment “that

---

<sup>28</sup>The phrase ‘numbing grossness’ is borrowed from Strawson (1966, p. 137).

consciousness [...] is a myth" (p. 154). Consciousness is something that we cannot get rid of without ceasing to be.<sup>29</sup>

Ryle faces a number of dilemmas: on the one hand, he's talking about the dispositions of *human* beings; he clearly states: "when we characterize as witty or tactful some piece of overt behaviour, we are not considering only the muscular movements which we witness. A parrot might have made the same remark in the same situation without our crediting it with a sense of humour." This implies that the kinds of dispositions that Ryle is talking about *presuppose* not just consciousness, which the parrot presumably has, but self-consciousness that persons have. On the other hand, Ryle insists: "to admit, as we must, that there may be no visible or audible difference between a tactful or witty act and tactless or humourless one is not to admit that the difference is constituted by the performance or non-performance of some extra secret acts" (p. 33). There is nothing extra over and above the behavioural dispositions. Rather, our understanding (or, should I say, our overt behavioural dispositions towards?) the complex dispositions involved in each utterance enables us to attribute wit (or, rather, to behaviourally dispose dispositions to the complex dispositions) to a person and not to a parrot. Adhering to Ryle's dispositional behaviourism we end up, not only with a category mistake, but with nonsense.

Later on in his book, Ryle states, in "considering specifically human behaviour – behaviour, that is, which is unachieved by animals, infants, and idiots – we should for several reasons notice the fact that some sorts of actions are in one way or another concerned with [...] other actions [...] the performance of the former involves the thought of the latter [...] my performance of my act involves the thought of your performance of yours" (p. 182). On the one hand, he distinguishes between human behaviour and that of animals (infants and idiots?), which implies that human behaviour presupposes self-consciousness and not simply consciousness. On the other, he claims that the notions of consciousness, self-consciousness, and the mental are "logical muddles" (149). Ryle wants to have his cake and eat it too.

Ryle distinguishes between what he calls "single-track dispositions" and dispositions "which are indefinitely heterogeneous." He refers to a vase's dispositions of fragility and shattering into fragments as "single-track dispositions, the actualisations of which are nearly uniform." But intelligent capacities of human beings, such as Jane Austen (to whom he refers as an example to illustrate his view (p. 44)), are multi-track dispositions, whose "actualisations [...] can take a wide and perhaps unlimited variety of shapes." These are dispositions "the exercise of which are indefinitely heterogeneous" (pp. 43–44). On Ryle's view the difference between a vase and Jane Austen is simply a matter of degree in terms of dispositions, not of *kind*. Even the dispositions themselves are not different in kind, but only in degrees of complexity, heterogeneity, and variety of shape.

If there is nothing but behavioural dispositions that we observe in ourselves and in others, as Broad (1925) pointed out, how did Ryle form any conception of intelligence, or the exercise of intelligence (p. 44) or intelligent capacities? How did he come to suppose that there are such mental states and mental capacities, and then launch an assault on them?

There is no escaping from the fact that Ryle's attributions of behavioural dispositions to persons *presuppose* a conception of mind and self-consciousness, unlike a vase's dispositions of fragility which don't; they presuppose an understanding of perception, reasoning, thinking as a condition of the possibility and intelligibility of that conception. Without that conception, Ryle's attempt to recast or

---

<sup>29</sup>These concerns apply to recent equally astonishingly reckless and fallacious claims that consciousness is an illusion, since illusions *presuppose* consciousness. Even if the proponents of such claims are in a state of permanent illusion, their pitiable condition still *presupposes* consciousness.

describe agents acting, observing, reasoning, thinking in terms of dispositions to behave would be unintelligible, and therefore self-defeating. If Ryle tries even to just describe (let alone explain) the notion of behavioural disposition, he is necessarily driven back to such *a priori* understanding, and so into an inescapable vicious circularity. No attempt to sever these implications of the central conception of behavioural dispositions can succeed.

The difference between an intelligent reasoner, such as Jane Austen, and a vase is not simply a matter of degree, nor is it simply in terms of indefinite heterogeneous multi-track dispositions, as Ryle states; the difference is a difference in *kind*. A person is a subject whose acts or actions, whose behaviour, can be *imputed* to him or her; an object, such as a vase, is that to which nothing can be *imputed*. A person is a subject in possession of freedom and responsibility, an inanimate object is not. But Ryle declares that the will and its volitions, and by implication freedom, are mythical, and aims to “discuss [such a doctrine] destructively” (the only other method Ryle knows, after abuse), and “emancipate these ideas from bondage to an absurd hypothesis” (pp. 61–62). A person can act, not simply in accordance with scientific laws, not even *in accordance with* principles and values, but from the *recognition* of such principles and values. An object, such as a vase, moves (not acts) in accordance with scientific laws when it’s pushed or pulled, and has no say in the matter. There’s a difference in *kind* between personhood and objecthood which demonstrates Ryle’s category mistake.

As Descartes writes, to say that a person “can exist although no mind exists is just as silly as saying that a person qualified in architecture can exist although no architect exists (provided, that is, the word ‘mind’ is taken in the ordinary sense, in the way in which I explained I was taking it)” — namely, a “thinking thing which in common usage is termed a ‘mind’” (*Seventh Set of Objections with Replies* AT VII 558, and AT VII 525 respectively). It’s clear that the word ‘mind’ is taken by Descartes in the ordinary common usage, yet Ryle accuses Descartes of distorting “ordinary language” by “misunderstanding how ordinary language functions.” But it now transpires that it’s Ryle who misuses ordinary common usage and distorts ordinary language, violating the logic of concepts.

The kinds of dispositions that Ryle attributes to persons, and the phrases and metaphors in which he is caught up and which he uses to describe them, may not *verbally* presuppose the existence of minds and mental states, but they do in reality.

### Ryle’s own ghost in the machine

Let’s now turn to Ryle’s “plea [that he is not] denying well-known facts about the mental life of human beings;” rather, his aim is to “explode the myth” (p. 17): the “special use of ‘mental’ and ‘mind’ in which they signify what is done ‘in one’s head’ cannot be used as evidence for the dogma of the ghost in the machine. It is nothing but a contagion from that dogma” (p. 35).

It’s hard to understand what Ryle means by the “well-known facts about the mental life of human beings” since ‘the mental life’ for him is nothing over and above a bundle of indefinite multi-track behavioural dispositions; either he presupposes the notion of the mental as it is ordinarily used in order for his claim to make any sense, or he contradicts his behaviourism. Ryle caricatures Descartes’ dualism as being committed to the dogma of the ghost in the machine. What Ryle misses is that Descartes preempts and rebuts any such commitment in at least two famous passages.

**Descartes’ first rebuttal.** In the best-known passage, Descartes says: “I am not merely present in my body as a [pilot] is present in a ship” (*Sixth Meditation* AT VII.81). Rather, I am substantially united with my body: “I and the body form a unit.” Descartes argues for the substantial union between the mind and the human body. What reasons does Descartes have for defending such a union? His argument is: “If this were not so,” if there were no substantial union between the mind and the human body, “I [...] would not feel pain when the body was hurt, but would perceive the damage purely by the intellect, just as a sailor perceives by sight if anything in his ship is broken. Similarly, when the body needed food or drink, I should have an explicit understanding of the fact, instead of having [...] sensations of hunger and thirst. For these sensations of hunger, thirst, pain and so on are [...] confused modes of thinking which *arise from the union* and, as it were, intermingling of the mind with the body”

(AT VII 81; italics added). 'Confused' from its Latin etymology '*confusus*' from '*confundere*' means mixed or mingled — the mixing of reason and the senses or the union of mind and the body. The term 'confused' in 'confused sensory perceptions' or 'confused modes of thinking' does not mean bewildered, or perplexed, or involving inherent confusion as some commentators think (for examples, see Alanen, 1996, pp. 3–20). It means mixed, not distinct, because they 'arise from the union' (AT VII 81). There is no inconsistency in saying that a sensory perception is clear and confused, since the latter is contrasted with distinctness, while the former is contrasted with obscurity. However, whatever is distinct must also be clear and must be true.

Armstrong (1968), unlike Ryle, does refer to Descartes' example of the pilot and the ship, but objects: "the pilot is not only causally related to his vessel, acting on it and being acted by it, but he is also spatially inside the vessel" (p. 25). The central point that Armstrong missed, however, is not whether there is a spatial relation, or whether there is causal interaction. Rather, Descartes' example demonstrates that the pilot and the ship are *not* substantially united, nor do they need to be in order to interact; they can causally interact without being united. If the ship is damaged, the pilot feels no pain; he simply has an understanding of the fact of the damage, gained through seeing the damage on the ship. But this is not the case with conscious beings; when their bodies are damaged they feel pain; they don't just have an understanding of the damage.

As Descartes replies to Gassendi: "when you try to compare the intermingling of mind and [the] body with the intermingling of two bodies, it is enough for me to reply that we should not set up any comparison between such things because they are quite different in kind" (*Fifth Set of Replies* AT VII.390), though they are both subject to the single *Causal Principle* and its neutrality as to the nature of the relata. He didn't make the unfounded assumption that causal relata are homogeneous, or that causation is a physical relation. The neutrality of the *Principle* bridges the metaphysical distinctness of the two substances (see Christofidou, 2019).

Why are the two sorts of interaction quite different in kind? When the problem is raised in this way, it becomes clear that there is an underlying thesis that needs to be brought out. It's enough for Descartes to reply to Gassendi as he does, because there is a *fact of the matter* regarding the difference in kind between any other causal interactions in the world and mind-body, body-mind bidirectional causal interactions: what is explicitly unique and marks that 'difference in kind' is a metaphysical necessity, the *presuppositionality*, of the substantial union, without which a person would not be a person, would not be a human being — the union is "a true mode of union, as everyone agrees, though nobody explains what this amounts to" as Descartes writes to Regius (January 1642 AT III 493 & (508); CSMK:206 & 209).

What it amounts to, and what nobody has explained – then or now – is not only the truth of the substantial union, but its *uniqueness*. It is unique because innumerable causal interactions – body-body interactions – occur in the world without a presupposition of a union — without presupposing a principle of unity among them. This is the principle that Hume (1978) insightfully captures: "there is no question in philosophy more abstruse than that concerning [...] the nature of the uniting principle, which constitutes a person" (I.iv.ii). The bidirectional causal interaction between mind and the human body does not constitute the union but presupposes their union. A substantial union is presupposed by *what it is to be* a human being, a person, a self-conscious being, or by *what it is to be* a sentient conscious creature. The unity of what it is to be a person, unlike that of a machine, a computer, or a robot, is "ordained by nature" (*Optics* AT VI.130). The machine is not "a true entity in itself (*ens per se*) but "an accidental entity (*ens per accidens*)."

Mind-body bidirectional causal interactions are different in kind, as Descartes rightly stresses, because they presuppose the union. Any explanations of body-body causal interactions cannot therefore

be used either against or in defence of mind–body bidirectional causal interactions, because these interactions could not occur without the mind-body substantial union. The reason for this constraint, and the answer to the question above, is that without the substantial union we would *not* have experiences, sensations, pains, feelings, or sense perceptions. Nor would the mind ‘incline its will’ to action (Letter for [Arnauld] 29 July 1648 AT V.222; CSMK:357).

Without the presupposition and uniqueness of the substantial union, we would be detachedly aware of causal effects, like pilots in ships (AT VII 81), having simply ‘an explicit understanding of the facts.’ But that’s all contradicted by the irreducible and undeniable facts of conscious and self-conscious awareness. The substantial union is the only way to understand *how* we are, and *why* we feel so intimately bound up with our bodies.

Metaphysically, however, the substantial union does not follow from our experience or first-person awareness of interaction: it is presupposed by the interaction, a presupposition that can be clearly grasped by the intellect (see Christofidou, 2019). It’s an objective fact that, if there is sentience or consciousness and not only self-consciousness, then there must be a substantial union. There is, it seems to me, a clear parallel between the metaphysics of the substantial union and the metaphysics of the unity of consciousness.

Descartes, insightfully, turns past and present theses of mind-body interaction on their heads: there is no more profound unity than that of mind-body union presupposed by mind-body interactions. Whenever and in whatever case the mind is involved, there must be a true unity; in the case of rational free active beings, all the mind’s acts presuppose the principle of unity of self-consciousness, which in turn presupposes an irreducible *I* or self. Metaphysically, the substantial union is the only way to understand *what* we are as persons: substantially embodied rational agents who take epistemic and moral responsibility for our acts and actions.

**Descartes’ second rebuttal.** The second famous passage in which Descartes defends the true union of the mind and the human body, is found in his reply to Arnauld’s objection. Arnauld says: “It seems [...] that the [real distinction] argument proves too much, and takes us back to the Platonic view (which M. Descartes nonetheless rejects) that nothing corporeal belongs to our essence, so that man is merely a rational soul and the body merely a vehicle for the soul — a view which gives rise to the definition of man as ‘a soul which makes use of a body’” (*Fourth Set of Objections* AT VII.203).

Descartes replies by referring to his *Real Distinction Argument* in the *Sixth Meditation* (AT VII.78) arguing that it’s not clear “why this argument ‘proves too much’. For the fact that one thing *can* be separated from another [...] is the very least that can be asserted in order to establish that there is a real distinction between the two” (emphasis added). He continues: “I thought I was very careful to guard against anyone inferring from this that man was simply ‘a soul which makes use of a body.’ For in the *Sixth Meditation*, where I dealt with the distinction between the mind and the body, I also proved at the same time that the mind is substantially united with the body. And the arguments which I used to prove this are as strong as any I can remember ever having read” (*Fourth Set of Replies* AT VII.228).

It’s evident that for Descartes a person is neither a construct or a neurobiological particular, nor a disembodied mind or ego, despite misattributions over the centuries. Nor is it “a ghost mysteriously ensconced in a machine,” contrary to Ryle’s careless claim, as is shown from Descartes’ denial of Arnauld’s suggestion that a person is simply a mind that happens to have a body as a vehicle (AT VII 203).<sup>30</sup> This is also shown by Descartes’ affirmation that a person is an irreducible, unanalysable true entity — neither analysable nor reducible to either a mind or a body (*Comments on a Certain Broadsheet* AT VIII B 351). A person, for Descartes, is “a true entity in itself (*ens per se*) not an accidental entity (*ens per accidens*)” (Letter to Regius January 1642 AT III.493; CSMK:206. See also the *Sixth*

---

<sup>30</sup>Brown (2013, p. 55) states that in his reply to Gassendi, Descartes claims: ‘the mind *uses* a body’ contrary to his response to Arnauld. But Descartes is concerned there with the mind’s ‘power of moving the body’ (AT VII.389), of interacting with it, not of using it.

*Meditation*, his letters to Arnauld, and to Princess Elizabeth.). Persons or human beings are embodied, natural, real irreducible individuals. For Descartes, the true substantial union of mind and the human body is constitutive of what it is to be a person — the union is “a true mode of union” not imposed or constructed by us.

The substantial union provides a framework in which to recognise not only the essential distinctness (not separate existence) between mind and the body, but also their interrelatedness; it enables us to articulate *that* distinctness within a single intelligible conception of nature. But the fact that the mind and the body are united “does not license the inference that they are one and the same” (*Sixth Set of Replies* AT VII 444), or that the mind is derivative or arises from the physical processes of the brain. It is a conceptual contradiction to think that ‘union’ is equivalent to ‘identity,’ or ‘dependence,’ or ‘derivativeness,’ or that by unifying matter and mind Descartes’ dualism is avoided.

In the *Second Meditation*, Descartes’ conception of the self and self-consciousness is epistemic and involves the self’s awareness of itself as an active real being indubitably existing in fact and in efficacy. In the *Sixth Meditation*, as we have seen, having gone through a number of arguments, and built on a number of stages of clear and distinct understanding, a conception based in an order corresponding to the actual truth of the matter (that is, metaphysically), self-consciousness encompasses the self’s awareness of itself as an active real being existing most crucially as an embodied irreducible human being, a person — an embodiment that presupposes the *union*.

The upshot is that Ryle displays not only a worrying ignorance of any of Descartes’ work, he also commits a two-fold category mistake: first, he fallaciously portrays *what it is to be* a person as a ghost inhabiting a machine without any connection to the body, and claiming that a person “lives through two collateral histories” an inner and an outer – which is nonsensical – when in fact a person is constituted by the mind-body substantial true union, a true entity in itself. Secondly, Ryle assumes that any mind-body interactions must “belong to one logical type or category,” to that of body-body causal interactions failing to realise that they actually belong to a unique category which presupposes a substantial union.

Thus, contrary to Ryle’s misattributions, for Descartes the substantial union between the human body and the mind is essential to *what it is to be* a person “since a human being [a person] without it [human body] is not a human being” (Letter to Regius January 1642 AT III (508); CSMK:209): a person is a unity *per se*.

Unlike Ryle (1949), for Descartes the human body is not another “parcel of matter, in the field of causes and effects” (p. 20); a parcel of matter could not be united with the mind, and although a parcel of matter is divisible, the human body “is a unity which is in a sense indivisible because of the arrangement of its organs” and the mutual dependence of its parts (*Passions of the Soul* 1:30 AT II 351). For Descartes “the human body has all the dispositions required to [be united with a] soul,” and without these dispositions it is not “strictly a human body” (Letter to Regius December 1641 AT III.461; CSMK:200), including “in itself all the dispositions required to preserve that union” (Letter to Mesland 9 February 1645 AT IV.166; CSMK:243). If the body has these dispositions “then short of a miracle it must be united to a soul” (AT III.461).

When discussing the human body, by ‘mechanical’ Descartes means organic, or of the organism, something of which Ryle seems ignorant, though he’s not alone in this. Reference to a machine’s interrelated parts is analogical, a way of explaining the interrelated organs of living beings whose unity (unlike that of machines) is not imposed by us (*Passions* I:30 AT XI.351), but “ordained by nature” (*Optics* AT VI.130). “It was in German Romanticism [Schlanger argues (1971:51)] that the contrast of the organic and the mechanical took on the value it still has, and the mechanical became coincident with the *inert*, the *lifeless*” (Des Chene, 2001, 68).

Ryle shows in one claim after another that he has not read Descartes' works, let alone made any attempt at understanding Descartes' commitments, despite claiming that his aim is to show that Descartes' dualism is false in principle — that *on its own terms* it is logically untenable. Ryle also attributes to Descartes the ideas of privileged access and introspection (16),<sup>31</sup> again in ignorance of Descartes' rejection of its usefulness, not least as a source of justification; in fact, Descartes is not concerned with justification, nor is he concerned with psychological certainty, both of which are consistent with falsehood. He is concerned with indubitability, with what cannot be doubted, and discovering what is real and true. And privileged access is not something that he ever considered, nor does it fit with his conception.

Ryle puts forward systematically misleading claims. He moves fallaciously from what he takes to be true to presenting such claims as being true. He falsely attributes theses to Descartes without any evidence. In contrast, Descartes, following his strict order of reasoning, states clearly that if there is something he does not know he does not argue for it (*Second Meditation* AT VII 27; see also **Kinds of distinction** in the second section above). The contrast is profound, pointing up the immense difference between the two thinkers in their respective approaches to proper philosophical enquiry and analysis.

In fact, to claim, as Ryle does, that the mind, or consciousness, or mental states are nothing over behavioural dispositions is to “misunderstand how ordinary language functions,” a “violation of the logic of concepts,” a distortion of ordinary language.” It is Ryle's behaviourism “that represents the facts of mental life as if they belonged to one logical type or category [behavioural dispositions], when they actually belong to another” (p. 17). Ryle commits a category mistake of numbing grossness: “The dogma is therefore a philosopher's myth” (p. 17), and that philosopher is Ryle.

The myth, the category mistakes, the dogma of the ghost in the machine are Ryle's own commitments which conflate the categories, eliminate the mind as a true natural entity, and thus obliterate all conscious creatures and all self-conscious persons. All this is achieved (if achievement is the right term) by speaking abusively, destructively, thriving on rhetoric, analogies, slogans, and presupposing his behaviourist thesis, thus begging the question against dualism by assuming what had to be proved.

Ryle's professed aim was to exorcise Descartes's dualism's “ghost in the machine.” Such an aim was not exactly difficult to accomplish since in Descartes' dualism there is no such ghost, and there is no such machine. This is all a fable, a myth conjured up by Ryle without any philosophical debate or engagement with the thesis under attack. It seems that “the assumptions against which I exhibit most heat” and “of which I have been a victim” (pp. 10–11), as Ryle states at the beginning of his book, are of his own making.

### A General Evaluation: Ryle's Highly Paradoxical Idea

The most devastating blow to Ryle's entire thesis, a *reductio ad absurdum*, which he didn't see, is his denial that the mind is a real natural entity. According to Ryle: “the logical geography of concepts is to reveal the logic of the propositions in which they are wielded, that is to say, to show with what other propositions they are consistent and inconsistent, what propositions follow from them and from what propositions they follow” (p. 10). But if there is no mind (or self), as Ryle insists, then no two thoughts, no two propositions can be connected and a third inferred, no propositions can follow from them, because there is nothing to connect the propositions. Consider: (T1) ‘it is thought: *P*’. (T2) ‘it is thought: *Q*,’ will never lead to any conclusion: (T3) ‘it is thought: *P* and *Q*,’ unless both propositions

---

<sup>31</sup>Ryle is not alone — almost every discussion in the Philosophy of Mind mistakenly attributes the notions of privileged access, introspection, and infallibility to Descartes. Badly and uncritically formed philosophical habits die very hard, if at all. What I find not only difficult to comprehend, but also indicative of a failure of intellectual integrity, is Kenny's attributions of any ill-conceived idea to Descartes, because Kenny, unlike many others, has published a study on Descartes. It's one thing to fail to understand a philosopher, it's quite another intentionally to misrepresent them (see Kenny, 1966, 1968, 1989).



are thought by the same (irreducible) mind or self (see Williams, 1978, p. 96). The same applies, *mutatis mutandis*, to any putative logically contradictory propositions, any disjunctive propositions, and so on.

Without the mind, no former thought can “involve the thought of [a] latter,” despite what Ryle claims, because there is nothing to connect the two thoughts. It would be unavailing for Ryle to retort that they are thought by the same person because a Rylean mindless person, as he insists, is simply a bundle of heterogeneous indefinite multi-track bundles of indefinite multi-track dispositions — and so *ad infinitum*. It would be profoundly nonsensical to claim that dispositions or bundles of dispositions can think, be in pain, or connect two bundles of dispositions and infer a third. Related to this is the question of how any text and notions can be conveyed in any way, let alone in any depth, in order to reveal not only the sense but also the creative effort undertaken.

Worst still, not only the connection or unity *between* thoughts or propositions, but also the connection or unity *within*, the compositionality of, a whole thought or proposition requires a mind or a self. Without the mind no complete thought, no reasoning, no judgement, no doubting, no understanding, no knowing (neither knowing *how*, nor knowing *that*) would be possible; this is because there would be nothing to connect the concepts in a proposition, nothing to connect the grammatical/logical categories of subject and predicate and form a well-structured, meaningful, and complete thought or proposition, such as ‘*a is F.*’ Only in so far as there is a mind or a self can a complete thought or proposition be formed and grasped. Without the mind or the self our entire system of reasoning, thinking, judging would collapse.

That’s what decisively undermines, not only Ryle’s thesis, but all reductionist and eliminative theories of the mind, all physicalist, neuronist, dispositional/power behaviourist theories, and non-physicalist property-dualist theses, since properties cannot think, judge, feel pain, or act, and shows that the requirement for a mind or a self, an irreducible subject of thoughts, is not a logical muddle as Ryle assumes. In fact, every sentence in Ryle’s book presupposes that there is a mind and a unity of self-consciousness, without which there would be no sentence, let alone a coherent sentence.

Ryle’s view generates a highly paradoxical idea of a human being, a person. The separateness, self-identity, and self-awareness of a person are all denied by Ryle. Indeed, it’s reasonable to argue that there is no place for the concept of an individual, or for the distinction, so important to our ordinary language and thought, between an individual and its dispositions, since there is no self or mind but only bundles of indefinitely heterogeneous multi-track behavioural dispositions, and hence no way of understanding the separateness and individuality of persons.

The consequence of this highly paradoxical and impoverished metaphysics is that it leaves the whole of morality, not in doubt, but impossible; it’s unclear whether the question ‘What ought *I* to do?’ has any meaning for Ryle, since there is no self and no *I*. According to Ryle, the concept of ‘I’ is “systematically elusive,” there is nothing for which it stands. It’s a shadow cast by grammar: “Like a shadow of one’s head, it will not wait to be jumped on. [...] It evades capture by lodging itself inside the very muscles of the pursuer. [...] Theorists have found themselves mocked in a similar way by the concept of ‘I’” (p. 178).

This, seems to me, is what Hume (1978) recognised and had the intellectual honesty to openly admit and lament in the *Appendix* to the *Treatise*: that with his bundle theory he had failed to account for the self and, given the constraints of his empiricism, had failed to account for “the nature of the uniting principle, which constitutes a person” (I.iv.ii). Consequently, none of his other theses could be sustained or make any sense (a topic I pursue elsewhere). And it’s precisely that recognition that prompts Hume to develop, in Book II of the *Treatise*, a positive account of the self so that his moral thesis could get off the ground. The impossibility of subjectless or mindless thoughts is not epistemic,

or psychological, or grammatical, but metaphysical. The very notion of thought itself presupposes the notion of a mind or self as a true entity.

The key point seems to be that self-conscious agents, by the very fact of self-consciousness, are capable of reflecting upon themselves, and in so doing to conceive alternatives to their thoughts, feelings, desires, and actions; they are thereby also capable of conceiving the mental and emotional states of others, and of responding to them in a non-impulsive way, but with sympathy and kindness. Conversely, it's precisely their ability to conceive of the mental states of others that enables them to carry out acts of betrayal, deception, and revenge. None of this would be possible if the states of the mind of others were inaccessible and private as Ryle insists. Nor would there be any ground for intrapersonal negative moral emotions such as guilt, shame, and resentment, nor for interpersonal positive moral emotions such as empathy, sympathy, and compassion.

It appears that Ryle in his derision, in his polemic and destructive approach, in replacing philosophical engagement by mere abuse, though claiming to be demolishing confusions and logical blunders is blinded to what is true, to what is a true requirement for any reasoning, theoretical and practical, for any sensing, feeling, for any complete proposition to be formed, to be connected with other propositions, to be entailed by other propositions, to be consistent or inconsistent with other propositions.

It is not that Ryle simply failed to “point to any flaw whatsoever in my arguments” (*Sixth Set of Replies* AT VII 446), as Descartes would have responded, but he failed to engage with *any* of Descartes' arguments. Ryle not only failed in his aim to show that Descartes' dualism is false in principle, but he failed to show that it is false at all. We have no reason to accept Ryle's polemic; indeed, even his polemic is self-refuting, it collapses on itself, because it cannot get off the ground without presupposing a true natural entity: the mind.

### Epilogue

In the Introduction to *The Concept of Mind*, Ryle writes: “Philosophy is the replacement of category-habits by category-disciplines, and if persuasions of conciliatory kinds ease the pains of relinquishing inveterate intellectual habits, they do not indeed reinforce the rigorous arguments, but they do weaken resistance to them” (p. 10).

What's most worrying is how Ryle's self-refuting thesis, his caricatures, unfounded attributions, hasty generalisations, and misleading claims have become inveterate philosophical habits perpetuated from one generation of thinkers to another.

Physicalist and non-physicalist philosophers, psychological and philosophical behaviourists, neo-dispositional/power behaviourists, reductionists, neuronists, eliminativists through intellectual habit continue to saddle Descartes with whatever ill-conceived theses they conjure up, and subsequently define their positions against this invention (see also Cottingham 1994, p. 2; Christofidou, 2001). Such habits, it seems, cannot be based, as Descartes argues, on “preconceived opinions of childhood” (*First Meditation* AT VII 17) even if “it is not easy for the mind to erase [...] false judgements [or preconceived opinions] from its memory” (*Principles* I, pp. 72–73). If so, on what is philosophical habit based? What is the epistemic and metaphysical significance of uncritically formed intellectual habits?

The notions of uncritically formed habits and preconceived opinions are related to what Descartes understands by philosophical prejudice.<sup>32</sup> A philosopher who (perhaps more than anyone, along with Socrates and Plato) aimed at liberating the mind from preconceived opinions and prevailing

---

<sup>32</sup>In contemporary philosophy, one such prejudice, as we have seen, is the physical causal-closure assumption. Another is the insistence that we adopt the peculiar “dogmatic rule that dualism is to be avoided *at all costs*” (Dennett, 1991, p. 37).

authoritarian doctrines, has been the subject of persistent misattributions as a result of philosophical prejudice and inveterate intellectual habits.

How is philosophical habit formed? Descartes suggests various ways; for example: “Men of learning [...] have got so used to elaborating their contrived doctrines that they have gradually come to believe them and to pass them off as true” (*Rule Two* AT X.362–3). These doctrines are passed on, and accepted by those who come afterward, partly because of the confidence with which they are articulated, and partly because of the authority with which they are presented, as we have seen in Ryle’s attacks which rely heavily on his confidence and the authority he assumes. In contrast, Descartes says: “the greatest pleasure I have taken in my studies has always come not from accepting the arguments of others but from discovering arguments by my own efforts” (*Rule Ten* AT X.403) through painstaking cross-examination. There is no doubt that Descartes’ dualism has been, through the centuries, rejected, attacked, denied, caricatured, scorned, assumed to be false, but not refuted.

Descartes’ general conception of philosophy is as an activity, an enquiry that would lead to what is real and true, and “the ultimate level of wisdom” (*Principles, Preface* ATIXB 5–9), rather than as a specific discipline or research programme. This conception of philosophy and the concern with the right method of engaging in philosophical enquiry are integral, not only to his metaphysics and epistemology and their implications for morality, but also to self-transformation and the formation of habits *critically* acquired — of virtues epistemic and moral. In Descartes’ hands the Socratic *elenctic* method of cross-examination is first and foremost self-administered.

The self, mind, and freedom stand or fall together. Descartes’ non-negotiable conception of freedom, its internal relation to reason, and its centrality to his entire metaphysical quest, is expressed in his letter to Queen Christina: freedom of spontaneity or of the higher degrees is “the greatest good [...] the supreme good [...] the noblest thing we can have” (20 November 1647 AT V.81–86; CSMK:324–326).

It seems to me that our deepest concerns and responsibility in the twenty-first century – morally, epistemically, and metaphysically – must be to uphold the irreducibility of the mind as a true real natural entity, and thereby save the authority of reason and the autonomy of freedom from being declared illusory and illegitimate by physicalism, neuronism, neo-dispositional/power behaviourism, and ill-formed philosophical habits.

### References

- Alanen, L. (1996). Reconsidering Descartes’ notion of the mind–body union. *Synthese*, 106, 3–20.
- Aristotle (1986). *Peri Psyche. Aristotle: On the Soul*. W.S. Hett (Trans.). The Loeb Classical Library, vol. VIII. Cambridge, MA: Harvard University Press.
- Armstrong, D. (1968). *A materialist theory of the mind*. London: Routledge & Kegan Paul.
- Baum, W. M. (2017). *Understanding behaviorism, Third Edition*. Malden, MA: Wiley.
- Broad, C.D. (1925). *The Mind and its place in nature*. London: Kegan Paul.
- Brown, D. (2013). Understanding interaction revisited. In S. Duncan & A. LoLordo (Eds.), *Debates in modern philosophy: Essential readings and contemporary responses* (pp. 54–64). New York: Routledge.
- Burgos, J. E., & Killeen, P. (2018). Suing for peace in the war against mentalism. *Perspectives on Behavior Science*, <https://doi.org/10.1007/s40614-018-0169-2>.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Christofidou, A. (2001). Descartes’ dualism: Correcting some misconceptions. *Journal of the History of Philosophy*, 34, 215–238.
- Christofidou, A. (2007). God, physicalism, and the totality of facts. *Philosophy*, 82, 515–542.

- Christofidou, A. (2009). Self and self-consciousness: Aristotelian ontology and Cartesian duality. *Philosophical Investigations*, 32, 134–162.
- Christofidou, A. (2016). *Self, reason, and freedom: A new light on Descartes' metaphysics*. New York: Routledge.
- Christofidou, A. (2019). Descartes: A metaphysical solution to the mind-body relation and the intellect's clear and distinct conception of the union. *Philosophy*, 94, 87–114.
- Christofidou, A. (under review). Descartes' cogito: the self's reality, freedom, and openness to the objective world. An earlier version was presented at a workshop in Ligerz, Switzerland, April 2017.
- Clarke, D. M. (2003). *Descartes's theory of mind*. Oxford: Oxford University Press.
- Coleman, S., & Alter, T. (in press). Panpsychism and Russellian monism. In W. Seager (Ed.), *The Routledge Handbook of Panpsychism*. New York: Routledge.
- Cottingham, J. (1994). Introduction: *Plus una vice agendum*: Cartesian metaphysics three and a half centuries on. In J. Cottingham (Ed.), *Reason, will, and sensation: Studies in Descartes' metaphysics* (pp. 1–16). Oxford: Oxford University Press.
- Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown & Co.
- Des Chene, D. (2001). *Spirits and clocks: Machine and organism in Descartes*. Ithaca, NY: Cornell University Press.
- Descartes, R. (1985). *The philosophical writings of Descartes, Vols. I and II* (John Cottingham, Robert Stoothoff, Dugald Murdoch, Trans.). Cambridge: Cambridge University Press.
- Descartes, R. *The philosophical writings of Descartes, Vol. III: The Correspondence* (John Cottingham, Robert Stoothoff, Dugald Murdoch, Anthony Kenny, trans.). Cambridge: Cambridge University Press.
- Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.
- Frege, G. (1948). Sense and reference. *Philosophical Review*, 57, 209–230.
- Gibb, S. (2015). Defending dualism. *Proceedings of the Aristotelian Society*, 115, 131–146.
- Goff, P. (2017a). *Consciousness and fundamental reality*. Oxford: Oxford University Press.
- Goff, P. (2017b). Are electrons conscious? OUPblog 13 August 2017. <https://blog.oup.com/2017/08/electrons-consciousness-philosophy/>
- Harzem, P. (2004). Behaviourism for new psychology: What was wrong with behaviourism and what is wrong with it now. *Behaviour and Philosophy*, 32, 5–12.
- Hempel, C. (1980). Comments on Goodman's *Ways of Worldmaking*. *Synthese*, 45, 193–200.
- Hume, D. (1978). *Treatise of human nature*. L.A. Selby-Bigge (Ed.), Second Edition. Revised by P.H. Nidditch. Oxford: Clarendon Press.
- Kenny, A. (1966). Cartesian privacy. In G. Pitcher (Ed.), *Wittgenstein: The Philosophical Investigations* (pp. 352–370). New York: Anchor.
- Kenny, A. (1968). *Descartes: A study of his philosophy*. New York: Random House.
- Kenny, A. (1989). *The metaphysics of mind*. Oxford: Clarendon.
- King, P. J. (2006). One man's meat is another man's person. In A. Batthyany & A. Elitzur (Eds.), *Mind and its place in the world: Non-reductionist approaches to the ontology of consciousness* (pp. 61–76). Frankfurt: Ontos Verlag.
- Martin, C. B. (2008). *The Mind in nature*. Oxford: Oxford University Press.
- Nagel, T. (2012). *Mind & cosmos: Why the materialist neo-Darwinian conception of nature is almost certainly false*. Oxford: Oxford University Press.
- O'Donohue, W., & Kitchener, R. (1999, Eds). *Handbook of behaviourism*. New York: Academic Press.
- Papineau, D. (2001). The rise of physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (pp. 3–36). Cambridge: Cambridge University Press.
- Place, U. T. (1999). Ryle's behaviorism. In W. O'Donohue, & R. Kitchener (Eds). *Handbook of behaviorism* (pp. 361–398). New York: Academic Press.
- Ross, W. D. (1930). *The right and the good*. Oxford: Clarendon.
- Ryle, G. (1949). *The Concept of Mind*. Middlesex: Peregrine Books.
- Solomon, R. C. (1993). *The passions: Emotions and the meaning of life, Revised Edition*. Indianapolis, IN: Hackett Publishing Company.
- Strawson, P. F. (1996). *The bounds of sense: An essay on Kant's 'Critique of Pure reason.'* London: Methuen & Co Ltd.
- Warnock, G. (1958). *English philosophy since 1900*. Oxford: Oxford University Press.

Williams, B. (1978). *Descartes: The project of pure enquiry*. Sussex: The Harvester Press.  
Wittgenstein, L. (1953). *Philosophical investigations* (G.E.M. Anscombe, Trans.). Oxford: Basil Blackwell.

**Author's Note**

Address correspondence to: Andrea Christofidou, email: [andrea.christofidou@philosophy.ox.ac.uk](mailto:andrea.christofidou@philosophy.ox.ac.uk)  
Postal address: Andrea Christofidou, Keble College, Oxford, OX1 3PG, and Worcester College, Oxford,  
OX1 2HB, U.K.