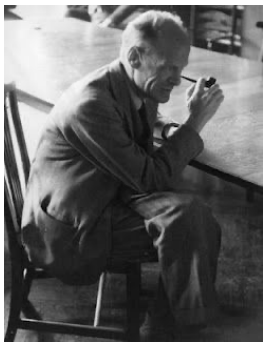
Descartes's Myth

Gilbert Ryle
From *The Concept of Mind*. 1949.
Edited and abridged by N. L. Engel-Hawbecker



1. The Official Doctrine

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. Most philosophers, psychologists, and religious teachers subscribe (with minor reservations) to its main articles. And although they admit certain theoretical difficulties in it, they tend to assume that these can be overcome without serious modifications being made to the architecture of the theory. It will be argued here that the central principles of the doctrine are unsound and conflict with the whole body of what we know about minds when we are not speculating about them.

The official doctrine, which hails chiefly from Descartes,[1] is something like this. With the doubtful exceptions of idiots and infants, every human being has both a body and a mind. Some would prefer to say that every human being *is* both a body and a mind. His body and his mind are ordinarily harnessed together, but after the death of the body his mind may continue to exist and function.

Human *bodies* are in space and are subject to the mechanical laws which govern all other bodies in space.[2] Bodily processes and states can be inspected by external observers. So a man's bodily life is as much a public affair as are the lives of animals, reptiles, trees, crystals, and planets.

But *minds* are not in space, nor are their operations subject to mechanical laws. The workings of one mind are not witnessable by other observers; its career is private. Only I can take direct notice of the states and processes of my own mind. 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, the second private. The events in the first history are events in the physical world, those in the second are events in the mental world. […]

It is customary to express this bifurcation of lives by saying that the things and events which belong to the physical world (including our bodies) are "external," while the workings of any mind are

[1] Rene Descartes (1596–1650). See the Canvas file "1.2 Descartes 1641 *Meditations*".

[2] Compare Descartes from page 5 of your second reading: "a body can't start up movements by itself and can move only through being moved by other things that bump into it. It seemed to me quite out of character for a body to be able to initiate movements… and I was amazed that certain bodies [namely, human ones] could do those things."

"internal." This antithesis of outer and inner is of course intended to be a metaphor, since minds (not being in space) could not be spatially inside anything else, or as having things going on spatially inside themselves. But relapses from this good intention are common, and theorists are found speculating how stimuli (the physical sources of which are yards or miles outside a person's skin) can generate mental responses inside his skull, or how decisions framed inside his cranium can set going movements of his extremities.

Even when "inner" and "outer" are construed as metaphors, the problem how a person's mind and body influence one another is notoriously charged with theoretical difficulties. What the mind wills, the legs, arms and the tongue execute; what affects the ear and the eye has something to do with what the mind perceives; grimaces and smiles betray the mind's moods, and punishing the body leads (it is hoped) to moral improvement. But the actual *transactions* between the episodes of the private [mental] history and those of the public [physical] history remain mysterious, since they can belong to neither series. They could not be reported among the happenings described in a person's autobiography of his "inner" life, but nor could they be reported among those described in someone else's biography of that person's "overt" life. They can be inspected neither by introspection nor by laboratory experiment. They are theoretical shuttlecocks which are forever being bandied from the physiologist back to the psychologist, and from the psychologist back to the physiologist.[3]

Underlying this partly metaphorical representation of the bifurcation of lives, there seems a more profound and philosophical assumption. It is assumed that there are two different *kinds of existence*. What exists or happens may have a physical existence, or it may have a mental existence. Somewhat as the faces of coins are either heads or tails, or somewhat as living creatures are either male or female, so (supposedly) some existing is physical existing, while other existing is mental existing. It is a necessary feature of what has physical existence that it is in space and time; it is a necessary feature of what has mental existence that it is in time but not in space. What has physical existence is composed of matter, or else is a function of matter; what has mental existence consists of consciousness, or else is a function of consciousness.

There is thus a polar opposition between mind and matter, an opposition which is often brought out as follows. Material objects are situated in a common field, known as "space," and what happens to one body is mechanically connected with what happens to other bodies in other parts of

[3] A shuttlecock is to badminton what a tennis ball is to tennis. In a volley, neither lands on one side of the court or the other. Ryle's picturesque metaphor suggests the same should be true for mind-body interactions.

space. But mental happenings occur in insulated fields, known as "minds," and there is (apart maybe from telepathy) no direct causal connection between what happens in one mind and what happens in another. Only through the medium of the public physical world can the mind of one person make a difference to the mind of another. A mind is its own *place*; in his inner life, each of us lives the life of a ghostly Robinson Crusoe.[4] People can see, hear, and jolt one another's bodies, but they are irremediably blind and deaf to the workings of one another's minds and inoperative upon them.

What sort of knowledge can be secured of the workings of a mind? On the one side, according to the official theory, a person has the best imaginable kind of "direct" knowledge of the workings of his own mind. Mental states and processes are (or are normally) conscious states and processes, and the consciousness which irradiates them can engender no illusions and leaves the door open for no doubts. A person's present thinkings, feelings, willings, perceivings, rememberings, and imaginings are intrinsically "phosphorescent"; their existence and their nature are inevitably betrayed to their owner. The inner life is a stream of consciousness of such a sort that it would be absurd to suggest that the mind whose life is that stream might be unaware of what is passing down it.

True, the evidence adduced recently by Freud[5] seems to show that there exist channels tributary to this stream, which run hidden from their owner. People are affected by impulses the existence of which they vigorously disavow; some of their thoughts differ from the thoughts which they acknowledge. […] However, holders of the official theory tend to maintain that in normal circumstances, a person must be directly and authentically [presented with] the present state and workings of his own mind.

Besides being currently supplied with these alleged immediate data of consciousness, a person is also generally supposed to be able to exercise from time to time a special kind of perception—namely inner perception, or introspection. He can take a (non-optical) "look" at what is passing in his mind. Not only can he view and scrutinize a flower through his sense of sight or listen to and discriminate the notes of a bell through his sense of hearing; he can also introspectively watch (without any bodily organ of sense) the current episodes of his inner life. This self-observation is also commonly supposed to be immune from illusion, confusion, or doubt. A mind's reports of its own affairs have a certainty superior to the best that is possessed by its reports of matters in the physical world. Sense-perceptions can be mistaken or confused, but consciousness and introspection cannot.

---

[4] In Daniel Defoe's novel *Robinson Crusoe* (1719), Crusoe is stranded on an island for 28 years.

[5] Sigmund Freud (1856–1939), Austrian neurologist and founder of psychoanalysis.

On the other side, one person has no direct access of any sort to the events of the inner life of another. He cannot do better than make problematic inferences from the observed behavior of the other person's body to the states of mind which, by analogy from his own conduct, he supposes to be signaled by that behavior. Direct access to the workings of a mind is the privilege of that mind itself; in default of such privileged access, the workings of one mind are inevitably hidden to everyone else. For the supposed arguments from bodily movements similar to their own to mental workings similar to their own would lack any possibility of observational corroboration. Not unnaturally, therefore, an adherent of the official theory finds it difficult to resist this consequence of his premises: he has no good reason to believe that there *do* exist minds other than his own. Even if he prefers to believe that to other human bodies there are harnessed minds not unlike his own, he cannot claim to be able to discover their individual characteristics, or the particular things that they undergo and do. Absolute solitude is the ineluctable destiny of the soul. Only our bodies can meet.

As a necessary corollary of this general scheme, there is implicitly prescribed a special way of construing our ordinary concepts of mental powers and operations. Those verbs, nouns, and adjectives, with which in ordinary life we describe people's wits, characters and higher-grade performances, must be construed as signifying special episodes in their secret histories, or else as signifying tendencies for such episodes to occur. When someone is described as "knowing," "believing," "guessing," "hoping," "dreading," "intending," "shirking," "designing," or "being amused," these verbs are supposed to denote the occurrence of specific changes in his hidden (to us) stream of consciousness. Only his own privileged access to this stream, in direct awareness and introspection, could provide authentic testimony that these mental-conduct verbs were correctly or incorrectly applied. […]

2. The Absurdity of the Official Doctrine

Such in outline is the official theory. With deliberate abusiveness, I shall call it "the dogma of the Ghost in the Machine." I hope to prove that it 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. Namely, it is a **category-mistake**. It represents the facts of mental life as if they belonged to one logical type or category (or range of types or categories), when they actually belong to another. The dogma is therefore a philosopher's myth [i.e., a self-imposed confusion, created by misusing terms]. In attempting to explode the myth, I shall probably be taken to deny well-known facts about the mental life of human beings, and my plea that I aim at doing nothing more than rectify the logic of mental-conduct concepts will probably be disallowed as mere subterfuge.

I must first indicate what is meant by the phrase "category-mistake." This I do in a series of illustrations.

1. A foreigner visiting Oxford or Cambridge for the first time is shown a number of colleges, libraries, playing fields, museums, scientific departments and administrative offices. He then asks "But where is the University? I have seen where the members of the Colleges live, where the Registrar works, where the scientists experiment and the rest. But I have not yet seen the University in which reside and work the members of your University." It has then to be explained to him that the University is not *another* collateral institution—some *ulterior counterpart* to the colleges, laboratories, and offices he has seen. Rather, the University is merely the *way* in which all that he has already seen is *organized*. When they are seen and when their co-ordination is understood, the University has been seen. His mistake lay in his innocent assumption that it was correct to speak of Christ Church, the Bodleian Library, the Ashmolean Museum *and* the University[6]—to speak, that is, as if "the University" stood for an extra member of the class of which the Church, Library, and Museum are members. He was mistakenly allocating the University to the same category as that to which the other institutions belong. […]
2. A foreigner watching his first game of cricket learns what are the functions of the bowlers, the batsmen, the fielders, the umpires, and the scorers. He then says "But there is no one left on the field to contribute the famous element of team-spirit! I see who does the bowling, the batting, and the wicket-keeping. But I do not see whose role it is to exercise team-spirit." Once more, it would have to be explained that […] team-spirit is not another cricketing-operation supplementary to all the other special tasks. It is, roughly, the *keenness with which* each of the special tasks is performed, and performing a task keenly is not performing two tasks [i.e., the task and the keenness]. Certainly, exhibiting team-spirit is not the same as bowling or catching, but nor is it a *third thing* such that we can say that the bowler first bowls and then exhibits team-spirit, or that a fielder is at a given moment either catching or displaying team-spirit.

The theoretically interesting category-mistakes are those made by people who are perfectly competent to apply concepts (at least in the situations with which they are familiar) but are still liable in their abstract thinking to allocate those concepts to logical types to which they do not belong. An instance of a mistake of this sort would be the following story.

[6] Christ Church is a college of the University of Oxford; the Bodleian Library and the Ashmolean Museum are also parts of the university.

3. A student of politics has learned the main differences between the British, the French and the American Constitutions, and has learned also the differences and connections between the Cabinet, Parliament, the various Ministries, the Judicature and the Church of England. But he still becomes embarrassed when asked questions about the connections between the Church of England, the Home Office[7] and the British Constitution. For while the Church and the Home Office are institutions, the British Constitution is not another institution in the same sense of that noun. So inter-institutional relations which [1] can be asserted or denied to hold between the Church and the Home Office [2] can*not* be asserted or denied to hold between either of them and the British Constitution. 'The British Constitution' is not a term of the same logical type as 'the Home Office' and 'the Church of England.'

In a partially similar way,

4. John Doe may be a relative, a friend, an enemy, or a stranger to Richard Roe. But he cannot be any of these things to the Average Taxpayer. He knows how to talk sense in certain sorts of discussions about the Average Taxpayer, but he is baffled to say why he could not come across him in the street as he can come across Richard Roe.

It is pertinent to our main subject to notice that, so long as the student of politics continues to think of the British Constitution as a counterpart to the other institutions, he will tend to describe it as a mysteriously [elusive] institution. And so long as John Doe continues to think of the Average Taxpayer as a fellow-citizen, he will tend to think of him as an elusive, insubstantial man. […]

My destructive purpose is to show that a family of radical category-mistakes is the source of the double-life theory. The representation of a person as a ghost mysteriously ensconced in a machine derives from [the following reasoning]. (P1) As is true, a person's thinking, feeling, and intending cannot be described solely in the idioms of physics, chemistry, and physiology. (P2) If so, then they must be described in counterpart idioms. Therefore, (C) just as the human body is a complex, organized unit, so too the human mind must be another complex, organized unit—though one made of a *different sort of stuff* and with a *different sort of structure*. Or again, as the human body (like any other parcel of matter) is a field of causes and effects, so too the mind must be *another field of causes and effects*—though not mechanical causes and effects (Heaven be praised).

[7] A government department of the United Kingdom, similar to the U.S. Department of State but also responsible for domestic security and immigration.

3. The Origin of the Category-Mistake

One of the chief intellectual origins of [that reasoning I suspect] to be this. When Galileo[8] showed that his methods of scientific discovery were competent to provide a mechanical theory which should cover every occupant of space, Descartes found in himself two conflicting motives. As a man of scientific genius, he could not but endorse the claims of mechanics. Yet as a religious and moral man, he could not accept […] the discouraging rider to those claims, namely that human nature differs only in degree of complexity from clockwork. The mental could not be just a variety of the mechanical.

He and subsequent philosophers naturally but erroneously availed themselves of the following escape-route. Since mental-conduct words do not signify mechanical processes, [the reasoning goes] they must signify *non*-mechanical processes. Since mechanical laws explain movements in space as the effects of other movements in space, [the reasoning goes] *other laws* must explain some of the *non*-spatial workings of minds as the effects of other *non*-spatial workings of minds. When we describe some human behaviors as intelligent and others as unintelligent, this difference *must* [the reasoning goes] be a difference in their *causes*. So while some movements of human tongues and limbs are the effects of mechanical causes, others must be the effects of *non*-mechanical causes—i.e., some issue from movements of particles of matter, others from workings of the mind.

The differences between the physical and the mental were thus represented as differences inside the common framework of the categories of "thing," "stuff," "attribute," "state," "process," "change," "cause" and "effect." Minds are [assumed to be] *things*, but different sorts of things from bodies; mental processes are [assumed to be] *causes and effects*, but different sorts of causes and effects from bodily movements. And so on. Somewhat as the foreigner expected the University to be an extra edifice (rather like a college but also considerably different), so the repudiators of mechanism represented minds as extra centers of causal processes (rather like machines but also considerably different from them). Their theory was a para-mechanical hypothesis.

That this assumption was at the heart of the doctrine is shown by the fact that there was from the beginning felt to be a major theoretical [task to explain] how minds can influence and be influenced by bodies. "How can a mental *process* (such as willing) cause spatial movements (like the movements of the tongue)? How can a physical change in the optic nerve have among its *effects* a mind's perception of a flash of light?" This notorious crux by itself shows the logical mold into which Descartes pressed his theory of the mind. It was the self-same mold into which he and Galileo set their mechanics. Still

[8] Galileo Galilei (1564–1642), Italian physicist and astronomer.

unwittingly adhering to the grammar of mechanics [i.e., its forms of speech], he tried to avert disaster by describing minds in what was merely an obverse vocabulary. The workings of minds had to be described by the mere negatives of the specific descriptions given to bodies: they are *not* in space, they are *not* motions, they are *not* modifications of matter, they are *not* accessible to public observation. Minds are not bits of clockwork; they are just bits of not-clockwork. Thus represented, minds are not merely ghosts harnessed to machines; they are themselves just spectral machines. Though the human body is an engine, it is not quite an ordinary engine, since some of its workings are governed by another engine inside it—this interior governor-engine being one of a very special sort. It is invisible, inaudible, and it has no size or weight. It cannot be taken to bits, and the laws it obeys are not those known to ordinary engineers. Nothing is known of how it governs the bodily engine. […]

It is a historical curiosity that it was not noticed that the entire argument was broken-backed. Theorists correctly assumed that any sane man could already recognize the differences between, say, rational and non-rational utterances, or between intentional and automatic behavior. […] Yet the explanation given [for these very differences] presupposed that, in principle, one person could never recognize the difference between the rational and the irrational utterances issuing from other human bodies, since he could never get access to the postulated immaterial "causes" of some of their utterances. Save for the exception of himself, he could never tell the difference between a man and a robot.[9] For example, it would have to be conceded that (for all that we can tell) the inner lives of persons who are classed as idiots or lunatics might in fact be as rational as those of anyone else ["on the inside"]. Perhaps only their overt behavior is disappointing. That is to say, perhaps "idiots" are not really idiotic, or "lunatics" lunatic [at least not "inside"]. Perhaps, too, some of those who are currently classed as sane are really idiots ["inside"]. According to the theory, external observers could never know how the overt behavior of others is correlated with their mental powers and processes. So they could never know *or even plausibly conjecture* whether their applications of mental-conduct concepts to other people were correct or incorrect. If so, it would also be irresponsible for a man to claim sanity or logical consistency *even for himself*, since he would be debarred from comparing his own performances with those of others. In short, our characterizations of people and their actions as "intelligent," "prudent," "virtuous," "stupid," "hypocritical," or "cowardly" could never have been made [if the official doctrine were correct]. So the problem of providing a special causal hypothesis to explain such diagnoses

[9] Again, from your reading by Descartes (page 8): "If I look out of the window and see men crossing the square, as I have just done, I say that I *see* the men themselves… yet do I see any more than hats and coats that could conceal robots? I *judge* that they are men."

would never have arisen. Put another way, the question, "How do persons differ from machines?" arose only because everyone *already knew* how to apply mental-conduct concepts *before* the new causal hypothesis was introduced. Therefore, that causal hypothesis could not capture why we use those terms. Nor, of course, has the causal hypothesis in any degree improved our handling of them. We still distinguish good from bad arithmetic, polite from impolite conduct, and fertile from infertile imaginations in the ways that Descartes himself distinguished them, before and after he speculated how those [terms might refer to hidden, hypothetical causes].

He had mistaken the logic of his problem. Instead of asking by what criteria intelligent behavior is actually distinguished from non-intelligent behavior, he asked "Given that the principle of mechanical causation does not tell us the difference, what other *causal* principle will tell it us?" He realized that the problem was not one of mechanics and *assumed* that it must therefore be one of some counterpart to mechanics. Not unnaturally psychology is often cast for just this role. […]

In unconscious reliance upon this dogma, theorists and laymen alike constantly construe the adjectives by which we characterize performances (as "ingenious," "wise," "methodical," "careful," "witty," etc.) as signaling the occurrence, in someone's hidden stream of consciousness, of special processes functioning as ghostly harbingers of the performances so characterized. They postulate an internal shadow-performance [i.e., in the mind] to be the real carrier of the intelligence ordinarily ascribed to the overt act. In this way, they think that they explain *what makes* the overt act a manifestation of intelligence—[namely, its being] an effect of a mental happening. Except they stop short before raising the next question: what makes the postulated mental happenings manifestations of intelligence?[10]

In opposition to this entire dogma, I argue that in describing the workings of a person's mind, we are not describing a second set of shadowy operations. […] The sense in which we "explain" his action is not that we infer to hidden causes, but that we subsume them under hypothetical and semi-hypothetical propositions. That is, the explanation is not of the [cause-and-effect] type "the glass broke

---

[10] This rhetorical question contains a famous argument. The Official Doctrine says that

(1) an action is intelligent only if it is caused by an act of the mind.

But acts of the mind (e.g. decisions) can themselves be intelligent or dumb, and it is implausible that a dumb mental act could produce intelligent behavior. So (1) is actually committed to saying

(2) an act is intelligent only if it is caused by an *intelligent* act of the mind.

But what makes that mental act intelligent? (2) tells us: it must have resulted from *another* intelligent mental act. But what made *that* second mental act intelligent? (2) tells us there must have been a third intelligent mental act. So if there are any intelligent acts at all, there would need to be an "infinite regress" of prior mental causes going back into the past. Unfortunately, no one has lived so long. So either there are no intelligent acts, or (more likely) the Official Doctrine is false.

because a stone hit it," but more nearly of the different type "the glass broke when the stone hit it, because it was brittle [and, as you know, brittle things break when hit like that]."[11]

[… So] when a person talks sense aloud, ties knots, feints, or sculpts, these actions which we witness are themselves what he does intelligently. The terms with which the physicist or physiologist would describe his actions do not exhaust those which would be [rightfully] used by his pupils or his teachers in appraising their logic, style, or technique. […] There is the one activity, but it allows and requires more than one kind of explanatory description. This is somewhat as there is no aerodynamical or physiological difference between the description of one bird as 'flying south' and of another as 'migrating', though there is a big biological difference between these descriptions. Like so, there need be no physical or physiological differences between the descriptions of one man as gabbling and another talking sense, though the rhetorical and logical differences are enormous.

The statement "the mind is its own place" (as theorists might construe it) is not true, for the mind is *not even a metaphorical* "place." On the contrary, the chessboard, the platform, the scholar's desk, the judge's bench, the lorry-driver's seat, the studio, the football field—these are where people work and play stupidly or intelligently. 'Mind' is not the *name* of another person, working or frolicking behind an impenetrable screen; it is not the *name* of another place where work is done or games are played; and it is not the *name* of another tool with which work is done, or another appliance with which games are played. [I maintain] that when we characterize people by mental predicates, we are not making untestable inferences to any ghostly processes occurring in streams of consciousness that we are debarred from visiting; rather, we are describing how those people conduct parts of their predominantly public behavior. […]

Overt intelligent performances are not *clues* to the workings of minds; they *are* those workings. […] I discover that there are other minds in understanding what other people say and do. In making sense of what you say, in appreciating your jokes, in unmasking your chess-stratagems, in following your arguments, and in hearing you pick holes in my arguments, I am not inferring to the workings of your mind, I am following them. Of course, I am not merely hearing the noises that you make, or merely seeing the movements that you perform. I am understanding what I hear and see. But this understanding is not inferring to hidden causes; it is appreciating how [i.e., the manner in which] the

[11] That is, we situate the person's action alongside familiar stereotypes (not necessarily positive or negative ones) or generic regularities in how similar people act in similar situations. For example, "They don't eat pork because they think it's sinful" is akin to, "She's going for a morning jog because she's a fitness-freak": neither explanation points to a *cause* that compels the action; each merely casts the relevant person as having a certain character, or as being a certain *sort* of person, who can only be expected to act in such ways. So recognized, their action no longer cries out for explanation.

operations are conducted. To find that most people have minds […] is simply to find that they are *able* and *prone* to do certain sorts of things. And this we do by witnessing the sorts of things they do. […]

Our inquiry is not into causes (and so not into hidden causes) but into capacities, skills, habits, liabilities, and bents. For example, we see a soldier score a bullseye. Was it luck or skill? [… To answer this, we] and he himself might need to take into account more than this one success. Namely, we should take into account his subsequent shots, his past record, his explanations or excuses, the advice he gave to his neighbor, and a host of other clues of various sorts. There is no one signal of a man's knowing how to shoot, but a modest assemblage of heterogeneous performances generally proves beyond reasonable doubt whether he knows how to shoot or not. [… For another example, a] drunkard at the chessboard makes the one move which upsets his opponent's plan of campaign. But spectators are sure that this was due not to cleverness but to luck, *if* they are sure that most of his moves made in this state break the rules of chess, or have no tactical connection with the position of the game, or that he would not be likely to repeat this move if the tactical situation were to recur, or that he would not applaud such a move if made by another player in a similar situation, or that he could not explain why he had done it or even describe the threat under which his King had been. [The spectators' hypothesis, then,] is not one of the occurrence or non-occurrence of ghostly processes, but one of the truth or falsehood of certain 'could' and 'would' statements. [In sum,] the difference between a normal person and a [mindless] idiot is not that the normal person really lives two lives while the idiot has only one. It is that the normal person can do a lot of things that the idiot cannot.[12]

4. Historical Note

It would be false to say that the official theory derives solely from Descartes' theories, or even from a more widespread anxiety about the implications of seventeenth century mechanics. […] Descartes was reformulating already prevalent theological doctrines of the soul in the new syntax of Galileo. […] It would also be false to say that the two-worlds myth did no theoretical good. Myths often do a lot of theoretical good, while they are still new. One benefit bestowed by the para-mechanical myth was that it partly superannuated the then prevalent para-political myth. Minds and their Faculties had previously been described by analogies with political superiors and political subordinates. The idioms used were those of ruling, obeying, collaborating and rebelling. [… At least] in anthropological and psychological theory, the new myth of hidden operations, impulses, and agencies was an improvement on the old myth of dictations, deferences, and disobediences.

[12] The last six paragraphs of this section were assembled and adapted from later in Ryle's book.