



Consciousness and Life

Gareth B. Matthews

Philosophy, Vol. 52, No. 199. (Jan., 1977), pp. 13-26.

Stable URL:

<http://links.jstor.org/sici?sici=0031-8191%28197701%2952%3A199%3C13%3ACAL%3E2.0.CO%3B2-W>

Philosophy is currently published by Cambridge University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/cup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

Consciousness and Life

GARETH B. MATTHEWS

In L. Frank Baum's story, *Ozma of Oz*, which is a sequel to Baum's much more famous story, *The Wonderful Wizard of Oz*, Dorothy and her companion come upon a wound-down mechanical man bearing a label on which are printed the following words:

Smith and Tinker's
Patent Double-Action, Extra-Responsive, Thought-Creating
Perfect-Talking
MECHANICAL MAN
Fitted with our Special Clock-Work Attachment
Thinks, Speaks, Acts, and Does Everything but Live
(*Ozma of Oz*, Chicago, 1907, p. 43)

As Dorothy and her companion are made to discover when they wind up this man ('Tik-Tok' is his name), he is indeed capable of doing all the things of which his label boasts—acting, speaking and even thinking. But as Tik-Tok himself insists, and no one in the story casts doubt on the matter, he is not alive.

Dorothy and her companion learn from the instructions that to make Tik-Tok think they are to wind under one of his arms, to make him talk they are to wind under the other, and to make him move about they are to wind in the middle of his back.

'Which shall we do first?' asks Dorothy. 'Wind up his think,' advises her companion. So she does. 'I don't see anything different,' says Dorothy after a moment's wait. 'Of course not,' explains her companion, 'he's just thinking; if you want to know *what* he's thinking, wind up his talk.' So Dorothy does. 'T-h-a-n-k y-o-u v-e-r-y m-u-c-h,' says Tik-Tok somewhat rustily.

Dorothy contrasts the case of Tik-Tok, who is incapable of living, with that of Nick Chopper, the much more famous tin man of *The Wonderful Wizard of Oz*. In *Ozma of Oz* Dorothy puts the contrast this way:

Once . . . I knew a man made out of tin, who was a woodman named Nick Chopper. But he was live as we are, 'cause he was born a real man, and got his tin body a little at a time—first a leg and then a finger and then an ear—for the reason that he had so many accidents with his axe, and cut himself up in a very careless manner (p. 42).

Dorothy's account is slightly inaccurate, as checked against the original story. Nick did not, according to his own account of the matter in *The Wonderful Wizard of Oz*, lose parts of his original body through carelessness, but rather because of a wicked witch who cast a spell on his axe.

But the really important point is that Nick started out as a creature of flesh and bones and ended up as a creature of tin by the gradual, piece-by-piece replacement of each limb and gross segment of his body.

The story of Nick's life is a modern variant of the ancient tale of the ship of Theseus, the boards of which were replaced one plank at a time. The puzzle with Nick, as with the ship, is to say whether the original entity persists through the gradual replacement of each of its parts and, if it does not, to say just at what point in the proceedings it ceases to exist.

Nick's story adds two new and conceptually significant features to the old tale of the ship of Theseus. First, the replacement pieces are, in his case, of a different material from the parts they replace. Each time he gets a tin replacement for a fleshly original. Just how and why the material might be important to the question of persistence through time is not easy to say. To gain some appreciation for how 'not easy' it is one need only have a look at Aristotle's tortured efforts in *Metaphysics* Z 10 to say whether, and if so, in what way, matter might belong to the form or essence or definition of a thing. The issues involved in this question become more pressing as we today become more and more clever at replacing parts of human bodies with fabricated substitutes.

The second new feature in Nick's story of gradual transformation is the fact that *he* tells the story as the story of *his* life. He seems able to remember being the creature of flesh and bones on whose axe the wicked witch cast a spell. In the case of persons, memory has long been a favourite candidate for a criterion of identity.

None of these considerations applies to Tik-Tok. He seems to have been fashioned in a tinker's shop. He is not the product of the gradual transformation of a living thing. He seems to be, as his label suggests, quite incapable of living.

One might want to draw a distinction between thinking and consciousness, to allow, for example, that a computer might be capable of thought, but not of consciousness. In the story, however, no such distinction is drawn; Tik-Tok is treated as a conscious being. So, if the story is coherent, neither

(1) Tik-Tok thinks

nor even

(2) Tik-Tok is conscious

entails

(3) Tik-Tok is alive.

That is, if the story is coherent there is nothing incoherent or self-contradictory about the claim that, while (1) and (2) are true, (3) is false.

In a paper called 'Robots: Machines or Artificially Created Life?' Hilary Putnam reports, as follows, an (as he says) 'ingenious argument', which he attributes to Paul Ziff:

Ziff wishes to show that it is false that [the robot] Oscar is conscious. He begins with the undoubted fact that if Oscar is not alive he cannot be conscious. Thus, given the semantical connection between 'alive' and 'conscious' in English, it is enough to show that Oscar is not *alive*. Now, Ziff argues, when we wish to tell whether or not something is alive, we do *not* go by its *behaviour*. Even if a thing looks like a flower, grows in my garden like a flower, etc., if I find upon taking it apart that it consists of gears and wheels and miniaturized furnaces and vacuum tubes and so on, I say 'What a clever mechanism', not 'What an unusual plant'. It is *structure*, not *behaviour* that determines whether or not something is alive; and it is a violation of the semantical rules of our language to say of anything that is clearly a mechanism that it is 'alive'.¹

I shall suppose that the talk of semantical rules at the end of this quotation can be expressed by saying that 'x is a mechanism' entails ' $\sim(x \text{ is alive})$ '. It would then be Ziff's claim (as reported by Putnam) that

(4) Tik-Tok is a mechanism

which, one might say, is proclaimed on Tik-Tok's label, entails

(3') Tik-Tok is not alive

and (3') in turn entails

(2') Tik-Tok is not conscious.

If we assume again that neither Putnam nor Ziff intends to make any crucial distinction between being conscious and thinking, we could as well say that (3') entails

(1') Tik-Tok does not think.

The upshot is that, according to Ziff, Tik-Tok's label would be self-contradictory.

For any of this to be at all plausible 'mechanism' must be understood in a rather hard-headed way, so that an expression such as 'the mechanisms of mind' would have to be metaphorical and so that in its literal use 'x is a mechanism' would entail that x operates on mechanical principles.

¹ 'Robots: Machines or Artificially Created Life?', *Journal of Philosophy* 61 (1964), 685-686; reprinted in Hilary Putnam, *Philosophical Papers* (Cambridge, 1975), Vol. 2, 402.

Ziff's reasoning (or anyway, the reasoning Putnam attributes to Ziff) is very traditional. One feels like saying that it is very Aristotelian. But there are good reasons for being cautious about saying that. For one thing, there seems not to be a good equivalent in Aristotle's Greek for the word 'conscious' and its cognates.² This apparent lack is not trivial; it has, in part, to do with the very separation of consciousness from life and bodily function that is under discussion.

For another thing, one might be worried about whether Aristotle had a concept of mechanism like ours. Evidence that he did is, perhaps, to be found in his discussion of mechanical toys as models for animal movement in *De Motu Animalium* (701b1) and as models for reproduction in *De Generatione Animalium* (734b10).

Then there are those hesitations about whether thinking, even among mortal things, requires a body (at, e.g., 403a6, 408b18, 413a6, 413b24)—hesitations resolved, perhaps, by Aristotle's doctrine that thinking requires an image, a phantasm (431a16, 431b2, 432a8, 432a13).

Certainly Aristotle is willing to suppose that eternal things think; in *Metaphysics* Λ 7 he tells us that the prime mover does. But then it also lives—not in the way you and I do, but in a way appropriate to itself. So if Aristotle were to conclude that Tik-Tok thinks, there is reason to suppose he would also conclude that Tik-Tok lives.

The discussions of toys I referred to earlier suggest that Aristotle would not be willing to call Tik-Tok, or Ziff's mechanical flower, alive. So perhaps it is fair to connect Aristotle with Ziff's reasoning. With respect to astronomy and theology, Ziff and Aristotle might disagree about what things think and therefore about what things are alive; but perhaps they are quite close on the connection between thinking and living and the separation between being a mechanism and either thinking or living.

By contrast, explicitly drawn contrast, Descartes sought to break the conceptual connection between being conscious and being alive. Here is a revealing comment from Descartes' *Reply to Objections V*:

. . . because probably men in the earliest times did not distinguish in us that principle in virtue of which we are nourished, grow, and perform all those operations which are common to us with the brutes apart from any thought, from that by which we think, they called both by the single name *soul* . . . But I, perceiving that the principle by which we are nourished is wholly distinct from that by means of which we think, have

² For helpful discussions of this question and the whole issue of how Aristotle's philosophy of mind bears on Cartesian ways of thinking see Charles Kahn, 'Sensation and Consciousness in Aristotle's Psychology', *Archiv für Geschichte der Philosophie* 48 (1966), 43–81, and especially Richard Sorabji, 'Body and Soul in Aristotle', *Philosophy* 49 (1974), 63–89.

declared that the name *soul* when used for both is equivocal . . . I consider the mind not as part of the soul, but as the whole of that soul which thinks.³

Descartes makes clear elsewhere that he wants to give an account of something's being alive in terms of its having in it warm blood. Here is one such passage:

. . . I would not wish to say that motion is the soul of brutes . . . I would prefer to say with the Bible (Deuteronomy, xii, 23, 'Only be sure that you eat not the blood; for the blood is the life; and you may not eat the life with the flesh') that blood is their soul; for blood is a fluid in rapid movement, and its more rarefied parts are called spirits. It is these which move the whole machine of the body as they flow from the arteries through the brain into the nerves and muscles.⁴

The mind, of course, has no blood in it (none, at least, in the relevant way, or perhaps in the relevant sense of 'in') and therefore is not alive. Yet the mind thinks, is conscious. Thus it does not follow from the fact that something is conscious that it is alive. Indeed, although the person made up of body and mind can presumably both be said to be alive and to be conscious, Descartes understands the person to be alive because the body is alive and conscious because the mind is conscious.

There is another way in which Descartes opposed the Ziffian argument. Descartes understands the corporeal soul to be a purely mechanical principle. He thus refuses to accept the alleged distinction between mechanisms and living organisms. The following quotation from Descartes' *Treatise of Man* expresses clearly Descartes' thesis that the living human body is a machine and operates on mechanical principles no different from those on which clocks operate:

I desire you to consider, further, that all the functions that I have attributed to this machine, such as [a] the digestion of food; [b] the beating of the heart and arteries; [c] the nourishment and growth of the members; [d] respiration; [e] waking and sleeping; [f] the reception by the external sense organs of light, sounds, smells, tastes, heat, and all other such qualities; [g] the imprinting of the ideas of these qualities in the organ of common sense and imagination; [h] the retention or imprint of these ideas in the memory; [i] the internal movements of the appetites and passions; and finally [j], the external movements of all the members that so properly follow both the actions of objects presented

³ Descartes, *The Philosophical Works*, Haldane and Ross (trans.), (Cambridge, 1931), Vol. 2, 210.

⁴ Letter to Buitendijck (1643), *Philosophical Letters*, A. Kenny (trans.), (Oxford, 1970), 146.

to the senses and the passions and impressions which are entailed in the memory—I desire you to consider, I say, that these functions imitate those of a real man as perfectly as possible and that they follow naturally in this machine entirely from the disposition of the organs—no more nor less than do the movements of a clock or other automaton, from the arrangement of its counterweights and wheels. Wherefore it is not necessary, on their account, to conceive of any vegetative or sensitive soul or any other principle of movement and life than its blood and its spirits, agitated by the heat of the fire which burns continually in its heart and which is of no other nature than all those fires that occur in inanimate bodies.⁵

Thus according to Descartes

(4) Tik-Tok is a mechanism

fails to entail either

(2') Tik-Tok is not conscious

or

(1') Tik-Tok does not think

both because it fails to entail

(3') Tik-Tok is not alive

and because (3') fails to entail (2'). In other words, Descartes rejected both a traditional connection and a traditional separation. The traditional connection he rejected is that between thinking, or being conscious, and living. He said that they do not really go together at all, except by a mysterious, divinely ordered coincidence. The traditional separation he rejected is the separation of living things and mechanisms; he said that living things are machines of a certain sort.⁶ The separation, like the connection, is traditional, though it was not universally accepted, even in the ancient world. Thales, according to Aristotle (*De Anima* 405a19), ascribed soul to magnets. We do not know for sure that he also said magnets are alive. But there is every reason to suppose he accepted the ancient commonplace that whatever has soul is alive. If so, and if a magnet counts as a mechanism, then Thales is a precursor of Descartes in denying the separation between living things and mechanisms.

⁵ Descartes, *Treatise of Man*, T. S. Hall (trans.), (Cambridge, Mass., 1972), 112–113.

⁶ What Descartes actually says in the last passage quoted is that the human body is *exactly like* a machine of a certain sort. Elsewhere (e.g. in the passage from his *Description of the Body* quoted by T. S. Hall on pp. 114–115 of his edition of the *Treatise of Man*) he simply says it is a machine.

Still, the ancient world did not, in general, follow Thales in this line of thought. And Descartes certainly gives every evidence that he thought himself something of a radical in conceiving the living human body as a machine. In fact he was so nervous about his radicalism and apprehensive of censure from officialdom that he disguised his claim in the *Treatise of Man* by making it explicitly a claim about a *model* of human physiology. (In other contexts he was more forthright.⁷)

As against Descartes, and in support of a somewhat Aristotelian position, Ziff claims that our language embodies a conceptual connection between being conscious and living and a conceptual incompatibility between being a mechanism and living. His claim is thus that our language remains resolutely anti-Cartesian on these matters.

Putnam, in his critique of the Ziffian argument, affirms the traditional connection between thinking and being alive. He speaks of 'the undoubted fact that if Oscar [the robot; we could substitute Tik-Tok, the Mechanical Man] is not alive he cannot be conscious' and of the 'semantical connection between "alive" and "conscious"'. But Putnam also shows some sympathy for Descartes' position in that he suspects the distinction between what is alive and what is mechanical is not clear-cut. It might be, he reasons, that some thinking thing, though alive, is yet also mechanical. The sympathy is only very limited. Descartes' claim was not really that you could make a mechanical man that would be alive; his claim was that living bodies *are* machines.

To make out the possibility that some robot might be alive Putnam maintains that criteria for life are of two quite different sorts—structural criteria and behavioural criteria. With plants it is, he says, structural criteria that are most important. With animals, or at least with human, or human-like, beings, behavioural criteria come into their own.

Crudely, Putnam's idea is this. Any robot with a programme that matched some reasonably good psychological theory of how people behave would thereby satisfy the behavioural criteria not only for being conscious but also for being alive. If it satisfied the *behavioural* criteria for life, but not the structural criteria, then so much the worse for the structural criteria in that case. It might, for all that, be alive.

What Ziff and Putnam call 'structural criteria' amount, I suppose, to some sort of biochemical definition of 'life'. By contrast the behavioural criteria would seem to belong to the family of functional criteria—digestion/absorption, growth, reproduction, self-motion, sensation/perception, and so on. It would be one thing to use two such basically different sets of criteria to get at roughly the same class of things. It is quite another thing to use one set of criteria for one group of putatively living things and

⁷ See previous footnote.

another set for quite a different group. To do that would be, it seems, tantamount to admitting that 'living' means something quite different when it applies to plants, say, from what it means as applied to human beings or human-being-like mechanisms.

Perhaps Putnam thinks that the expressions 'living' and 'alive' are fundamentally equivocal in just such a way. I myself think that the concept of life is complex and rather unclear in a number of ways. But Putnam's talk of two sets of criteria suggests that there are really two different concepts that can be rather neatly marked off from each other, and that our single English word 'alive' equivocally expresses now the one concept and now the other. I, for one, would have to be convinced that that is so.

What I want to concentrate on, however, is not what Putnam contests, but what he concedes, namely, that as our language now stands 'x is conscious' does indeed entail 'x is alive'. Speaking of the corresponding French and Latin words of his day Descartes in effect agrees with Putnam. But what Descartes says is that we had better alter the language so that this entailment no longer holds. As Descartes wanted to use the Latin and French words we render with 'thinks' and 'is conscious' and their cognates there would be no entailment from 'x is conscious' to 'x is alive'. I want to consider now why Descartes thought we ought to talk in this way.

In *Meditation II* Descartes concludes that he is a thinking thing, 'a thing which doubts, understands, affirms, denies, wills, refuses, which also imagines and feels', and that he is *not* 'a collection of members which we call the human body'. Moreover, he presents this double conclusion as a deliberate rejection of the traditional concept of soul.⁸

In the following passage from *Meditation II* Descartes encapsulates the traditional concept of soul as animator in confessing that he once held a belief he now rejects:

. . . I considered that I was nourished, that I walked, that I felt, and that I thought, and I referred all these actions to the soul . . .⁹

In these words one can recognize the traditional concept of soul as that whereby a human being is nourished, moves, perceives and thinks.

'But what am I', Descartes goes on to ask, 'now that I suppose there is a certain genius which is extremely powerful, and, if I may say so, malicious, who employs all his powers in deceiving me?'

This is a very peculiar question. Descartes does not ask, 'What do I *think* I am, now that I suppose such-and-such?' or 'What am I *justified* in thinking I am, now that I realize so-and-so?' He asks, 'What *am* I now

⁸ I am indebted to Professor G. E. M. Anscombe for making clear to me how explicit Descartes' concern with the traditional concept of soul is in *Meditation II*.

⁹ All translations from *Meditation II* are taken from *The Philosophical Works*, Haldane and Ross (trans.), (Cambridge, 1931), Vol. 1.

that I suppose I might be deceived in all sorts of diabolical ways?' Descartes seems to suppose that what he is, what, as we might put it, actually belongs to him, is going to be something impervious to the most extravagant doubts he can conjure up. The assumption seems to be that what he is, what really and truly belongs to him, is something that he cannot doubt belongs to him. Descartes thus seems to accept this principle:

(A) If I can doubt that x belongs to me, then x does not belong to me.

(A) looks at once too broad and too narrow to be a fundamental assumption. It looks too narrow in that it is restricted in its application to the person who enunciates it. (Or if we understand it as owned, so to speak, by Descartes, then we may be inclined to think of it as equivalent to

(A_d) If Descartes can doubt that x belongs to him, then x does not belong to him

which is even narrower.) Surely if (A)—or (A_d)—is acceptable, it is because it instantiates some more general principle, such as

(B) If x can doubt that y belongs to z , then y does not belong to z .

Or so one might reason.

(B), however, is unattractive in the extreme. From its universally quantified form one can conclude that everyone is infallible about everything.

Even (A) makes me infallible with respect to myself. Commentators have reached for this instead:

(C) If I can doubt that x belongs to me, then x does not belong to my essence.

(C) of course follows from (A), and there is anyway independent evidence that Descartes accepted something like it. But in the passage under consideration Descartes simply asks, 'What am I?'—not 'What am I essentially?' or 'What is my essence?' As we shall see, there is a special reason for his wanting to accept (A), as well, perhaps, as something more like (C).

Let us return to the narrative in *Meditation II*. Am I then a soul, as souls are traditionally conceived? That is, am I that being whereby my body is nourished, moves, perceives and thinks? Well, replies Descartes, 'If it is so that I have no body it is also true that I can neither walk nor take nourishment'. So I can doubt whether either being able to walk or to take nourishment belongs to me. Thus by (A) above neither does belong to me.

Gareth B. Matthews

Descartes moves on to perception, or sensation. 'But one cannot feel without a body', he reminds himself. Descartes then adds this curious remark:

... and besides I have thought I perceived many things during sleep that I recognized in my waking moments as not having been experienced at all.

What does Descartes suppose this remark relevant to?

His idea, I think, is this. There goes with the traditional concept of soul the idea that sensation and perception require a body. Consider, for example, this passage from Aristotle's *De Anima*:

In most cases it seems that none of the affections, whether active or passive, can exist apart from the body. This applies to anger, courage, desire and sensation generally, though possibly thinking is an exception. But if this too is a kind of imagination, or at least is dependent upon imagination [the alternative Aristotle seems to opt for], even this cannot exist apart from the body.¹⁰

It is not just that anger, courage, desire and sensation generally, *as it happens*, turn out not to be manageable without a body. It is rather that a proper account of what these things are will bring in the physiology of anger, courage, desire and sensation. One could say that, using 'sense' (or 'perceive') in a traditional way,

(5) x senses (perceives) something

entails

(6) x has a bodily organ of sensation (perception)

and thus

(7) x has a body (or is a body).

Descartes reasons that if I am going to use 'senses' and 'perceives' in this traditional way, I must say that sensation and perception do not belong to me, for I can doubt that I have a body at all and so I can doubt that I have bodily organs through which I might perceive or sense something. The remark about sleep is meant to give substance to this doubt.

A bit later on in *Meditation II* Descartes gives us a new meaning for '*sentire*' ('sense' or 'perceive'). In that new meaning the objectionable entailment does not hold. Here is what he says:

... I am the same who feels, that is to say, who perceives certain things, *as by organs of sense*, since in truth I see light, I hear noise, I feel heat.

¹⁰ 403a6-10, W. S. Hett translation, Loeb edition.

But it will be said that these phenomena are false and that I am dreaming. Let it be so; still it is at least quite certain that it *seems to me* that I see light, that I hear noise and that I feel heat. This cannot be false, *properly speaking* it is what in me is called feeling [*'sentire'*]; and used *in this precise sense* that is no other than thinking [or consciousness] (emphasis added).

This passage makes quite clear that Descartes thought of himself as recommending a way to use the terms for mental acts, or acts of consciousness. He wanted us to use them in such a way that the entailment from (5) to (6) and (7) will not hold—nor will the alleged entailment from (1), or (2), to (3).

In rejecting the traditional concept of soul and denying the traditional connection between consciousness and life Descartes does something much more interesting than simply isolating one function traditionally assigned to soul, namely, thinking, and supposing it to be independent of the rest. Instead he develops a new concept, consciousness, which includes thinking plus the 'inner part' (so to speak) of sensation and perception. Consciousness, conceived in Descartes' way, is the function of a self-transparent agent (the mind)—that is, of an agent whose acts and states are such that it cannot doubt that it has them.

The point is central. Let me have another try at stating it. In giving us the modern concept of mind Descartes does not simply isolate one accepted function of soul and say that a mind is something that does just that. Rather he divides up the functions of the soul in a new way and thinks about them in a new way. He supposes that a mind is a subject that performs a selection of soul-functions such that an entity cannot both perform one of those functions and also doubt that it is performing it. We must (so to speak) 'peel off' from seeing, hearing, tasting, etc., the *seeming* to see, to hear, to taste, etc., which is such that one cannot do that and also doubt that one is doing it. The mind is something that does all and only things of that self-revealing sort.

We can now see why Descartes accepts (A), and not just (C). One might accept (C) as an instantiation of this:

(D) If x can doubt that y belongs to z , then y does not belong to the essence of z .

Support for (D) could come from the idea that even in entertaining doubts about a thing (a case of thinking of the thing) one must know its essence, which knowledge in turn rules out doubting that y belongs to z when y belongs to the very essence of z .

The special reasons Descartes has for accepting (A) are very different. His idea is that there is a kind of entity, mind, which is self-transparent; for any act it performs and any state it is in, it cannot doubt that it performs

that act or is in that state. Furthermore, he, Descartes, is (he supposes) an entity of that kind.

I shall not try to say anything about the vexed question of how Descartes knows he is an entity of that kind, if indeed he does know it, or about what his best argument for proving it might be.¹¹ For my immediate purpose it is enough to say that Descartes is convinced that he is an entity of such a kind and that others have found it easy to think of themselves as entities of that kind.

The idea that there are entities of this peculiarly self-transparent kind has found wide acceptance in modern thought. Its apparent acceptability underlies the plausibility of, to pick one interesting illustration, Richard Rorty's article, 'Incorrigibility as the Mark of the Mental' (*Journal of Philosophy* 67 (1970), pp. 399–424, especially pp. 408–421). In this article Rorty considers in turn each of the following candidates for a criterion to distinguish what is mental from what is physical: intentionality, purposiveness, non-spatiality, introspectability, privacy as incommunicability, privacy as special access, privacy as unsharability, and incorrigibility. Rorty gives persuasive reasons for saying of each of these except the last, namely, incorrigibility, that it is unsatisfactory. Of special interest to him, and to us, is the problem of saying what 'intrinsic features . . . sensations and thought have in common' (p. 412) by virtue of which both sensations and thoughts count as something mental. It cannot be merely that they are soul functions, for on that basis nutrition and locomotion should also count. As I have already suggested, the idea that they have something fundamental in common, something that they do not share with a traditional soul function, such as locomotion, is largely a Cartesian innovation.

In the end Rorty also finds problems with incorrigibility as a mark of the mental (the problems have to do with the fact that avowals of one's present beliefs, desires, moods, emotions, intentions, etc., are not really incorrigible—though they are, he says, 'almost incorrigible' (pp. 419–420)). Rorty treats these problems as arising from *extending* the notion of the mental from thoughts and sensations to beliefs, desires, etc. It is not, I am afraid, historically true that the notion of the mental was first restricted to sensations and thoughts and then got extended to beliefs, desires, etc. Descartes clearly included beliefs, desires, emotions and intentions within the mental from the very first. Thus in *Meditation II* he says this:

But what then am I? A thing which thinks. What is a thing which thinks? It is a thing which doubts, understands, affirms, desires, wills, refuses, which also imagines and feels.

¹¹ Very helpful on these matters is G. E. M. Anscombe, 'The First Person', in *Mind and Language*, S. Guttenplan (ed.), (Oxford, 1975), 45–65.

The point is rather that Descartes was wrong in supposing that one cannot make mistakes about what one believes, desires, etc. Rorty's worry about whether first-person avowals of those things are truly incorrigible thus brings out a difficulty in the very concept of mind and the mental.

I have shifted rather naturally from talk of indubitability to talk of incorrigibility as a mark of the mental. Rorty gives a reason for preferring to speak of incorrigibility (p. 418) but, as he implicitly allows, the difference is small; historically it is clear, I think, that indubitability first played this role.

Subjects of consciousness, according to Descartes, get yoked (albeit temporarily) to living bodies. A human being, as we normally conceive one, is then a mysterious union of mind and body. But each being, so conceived, has both an 'inside' and an 'outside'. The 'inside' is a mind or subject of consciousness, whose acts and states are indubitable to itself. The 'outside' is a body known to the mind or 'inside' only indirectly through quite dubitable inferences from sense impressions.

The picture of human beings as having, in this way, both an 'inside' and an 'outside' is so commonplace, so (as it may seem to us) commonsensical that we find it hard to realize how strikingly modern it is. But to appreciate its modernity one need only cast about for statements of it earlier than Descartes. One does find interesting anticipations of it in Augustine, but not much earlier, and not much between the time of Augustine and that of Descartes. Here is a relevant passage from Augustine's *Contra Academicos*:

Do not assent to more than that you are convinced that it appears this way to you, and there is no deception; for I do not see how even an Academic can refute a man who says, 'I know that this looks white to me; I know that this sound pleases me; I know that this smells agreeable to me; I know that this tastes sweet to me; I know that this feels cold to me'. . . . If I am relishing the taste of something, who would be so shameless as to say to me, 'Perhaps you are not tasting, but this is a dream'? Would I stop? Why tasting would give me pleasure even in a dream. Wherefore, no semblance to falsity can confuse what I have said I know (Bk 3, sec. 11, chap. 26).

Just a little earlier on in the *Contra Academicos*, in chapter 25, Augustine suggests that one give the name 'world' to whatever impressions one has. Now, he announces triumphantly, this world exists and has whatever character it seems to have. No Academic sceptic can shake my incorrigible knowledge of my own inner world. Augustine's idea in *Contra Academicos* of an 'inner world', impervious to sceptical doubt, is Descartes' idea of the contents of one's mind or consciousness.

Gareth B. Matthews

Given this Cartesian starting point it becomes a central problem—one familiar to all modern students of philosophy—to justify one's belief that there is also an 'outer world' and that this outer world contains living organisms, including, it may be hoped, a living organism that one can identify as one's own body. Thus on Descartes' way of thinking and speaking there is no essential connection between being conscious and being alive, or even between being conscious and having a living body.

If Tik-Tok can carry on an intelligent conversation, there is no good Cartesian reason why he should not be said to be conscious on a basis analogous to that on which you or I might be said to be conscious. He could be a mind in mysterious union with a machine; his machine differs from your machine and mine only in that its physical principle of action is a heavy spring, rather than warm blood.

Suppose the problems of a Cartesian philosophy of mind are in fact as intractable as many people now suppose them to be. When one realizes that these problems arise from a rather deliberate decision by Descartes and his followers to conceive and talk about perception and thinking in a new way, coupled with a deliberate effort to wipe out all suggestion of conceptual incompatibility between being a machine and being alive, it becomes appropriate to ask whether it is a good idea to acquiesce in Descartes' proposal. A full answer to that question would require that one examine the coherence, and, if it is coherent, the usefulness of the idea of a thing whose acts and states are indubitable to itself. But short of giving that full answer one can at least gain perspective on the question by reminding oneself that there is a way of conceiving and discussing these matters, well established in philosophical tradition, in which Tik-Tok's label becomes, as no doubt Frank Baum intended that it should be, a conceptual joke.¹²

University of Massachusetts, Amherst

¹² Except for minor alterations the text of this article constitutes the third in a series of eight lectures given under the title, 'The Concept of Soul', in the Divinity School of Cambridge University in the Lent Term, 1976. I am indebted to Professor D.M. MacKinnon, to the other administrators of the Burney Fund and to the members of my wonderfully interdisciplinary audience for the opportunity to develop my ideas on the relevance of the traditional concept of soul to issues of contemporary concern.