the philosophical forum
A QUARTERLY

a philosophical quarterly published with the cooperation of the Department of Philosophy of Boston University

VOLUME VII, Nos. 3-4 SPRING-SUMMER 1976*

ARTICLES

DISCUSSION

*Erratum: Vol. VII, no. 2 was dated Spring 1976. It should have been Winter, 1975-76.

MECHANISM AND THE GOOD

JOSÉ A. BENARDETE

I

How a functionalist theory of value that is at once mechanistic and non-naturalistic (non-naturalistic in G. E. Moore's sense of the word and mechanistic in a suitable sense of the term to be explained) can be required by the very subject-matter of the life sciences, I propose to show in the following discussion, prompted in large measure by puzzlement over a single, fairly anomalous sentence in Descartes. That a mechanistic account of the world cannot escape being inherently normative when it comes to addressing itself specifically to the study of animals and perhaps even plants, is then the thesis I shall undertake to defend, showing in particular how the naturalistic fallacy can be avoided. Cartesian mechanism will be seen to afford the most natural mode of entry into the argument.

"The rules of mechanics . . . are the same as those of nature," writes Descartes (parenthetically) in the fifth part of the *Discourse on Method* after he has succeeded in showing in some detail how a purely mechanistic account can be given of the physiological processes in the human body. This mechanistic physiology is offered as being a mere fragment of a more comprehensive mechanistic account of the material world at large which Descartes feels prepared to sketch at least in outline. We are invited first to join in a piece of speculative cosmogony, leaving the present world and imagining another where there is only a sheer chaos of agitated matter "as confused as any the poets could feign". Relying solely on "the laws of nature", i.e. the rules of mechanics, Descartes professes himself able to show how eventually some of this agitated matter must come to compose an earth, some to compose planets and comets and some a sun and fixed stars. He then proceeds to "speak particularly of the earth" and how the rules of mechanics can explain the formation of mountains and rivers, the production of metals in the mines and even the growth of plants in the fields. With the mention of plants the question of

the origin and nature of life comes tacitly to the fore, and Descartes makes fairly plain his conviction (allowing for a certain nervous reticence brought on by the condemnation of Galileo) that a purely mechanistic explanation of life (its origin and nature) lies well within our reach, though he seems to suggest that animals might be somewhat more difficult to handle than plants. Even as regards man he has no doubt that mechanism can account for all bodily movements, both external and internal, with the decisive exception of those that result from the process of thinking, e.g. the movements of the lips in speech.

Descartes is not content, however, merely to project his mechanistic program in general terms; he is eager to show how very feasible it is, how indeed it can be executed on the level of precise detail, and availing himself especially of Harvey's recent discovery of the circulation of the blood he presents "the explanation of the movement of the heart and arteries" as a case-study of one of the bodily "functions" that will enable us "to judge easily what we ought to think of all the others." As for the "explanation" itself, it is pretty much the sort of thing that one expects to find today in any elementary textbook of physiology but there are certain incidental features of the account that repay examination, readers of Descartes' account being urged at the outset to have "cut up before their eyes the heart of some large animal that has lungs." One uninterrupted stretch of explanation I propose to quote at some length with the aim of soaking up its mechanistic flavor and acquiring a feel for its explanatory texture. Breaking then abruptly into his narrative, we read (in the Haldane and Ross translation):

> I should also wish that the eleven little membranes which, like so many doors, open and shut the four entrances which are in these two cavities, should be carefully shown. There are of these three at the entrance of the *vena cava*, where they are so arranged that they can in nowise prevent the blood which it contains from flowing into the right cavity of the heart and yet exactly prevent its issuing out; there are three at the entrance to the arterial vein, which, being arranged quite the other way, easily allow the blood which is in this cavity to pass into the lungs, but not that which is already in the lungs to return to this cavity. There are also two others at the entrance of the venous artery, which allow the blood in the lungs to flow toward the left cavity of the heart, but do not permit its return; and three at the entrance of the great artery, which allow the blood to flow from the heart, but prevent its return. There is then no cause to seek for any other reason for the number of these membranes, except that the opening of the situation where it is met with, may be conveniently closed with two membranes, while the others, being round, can be better closed with three. Further, I should have my readers consider that the grand artery and the arterial vein are much harder and firmer than are the venous artery and *vena cava* . . .

Embedded in this otherwise typical passage there is one sentence that

cannot but arrest our attention: "Et il n'est point besoin de chercher d'autre raison du nombre de ces peaux, sinon que l'ouverture de l'artère veineuse, étant en ovale à cause du lieu où elle se recontre, peut être commodément fermée avec deux, au lieu que les autres, étant rondes, le peuvent mieux être avec trois." A mystifying sentence, certainly. Descartes is here giving the "reason" why certain things are the way they are, and in each case the reason smacks very much more of Aristotle's formal and final 'causes' than of anything that a mechanist might be expected to countenance. Three things are to be explained, two thematically and one parenthetically. First, and almost incidentally, we learn the reason why it is that the opening of the venous artery is oval in shape, namely owing to (because of) the "place" where it is located, as if there were some natural fitness or inherent reasonableness for it to be occupying its 'proper' place. Second, we learn that the reason why there are precisely two valves at the opening to the venous artery is that by being oval it can the more readily be closed by two, as if nature were a careful, skilled workman who is to be understood as such. Third, and most striking of all, we learn that the reason why there are three valves at the other entrances is that each being round can be *better* closed by three than (presumably) by two or four.

If all of those 'explanations' ill-accord with what we are accustomed to demand from a strictly mechanistic account in terms of efficient causation the third is especially egregious when one considers the widely held view today that "scientific theory stands proudly and notoriously aloof from value judgments."[1] Any judgment as to what is good or bad, better or worse, being taken to be a value judgment, Descartes is here clearly engaged in making such a judgment and, not content with merely evaluating one state of affairs as being better than others, he explains why the better state of affairs obtains, on the ground that it is better. Surprising language from an agressive mechanist! That there are almost no other such 'lapses' in Descartes' account, and certainly none so glaring, renders it all the more puzzling that he should slip on this occasion. One can well understand how the historian of ideas might be tempted to find in the present case confirmation for the general hypothesis that no revolutionary thinker is ever as revolutionary as he fancies himself to be, being incapable of shaking off the accumulated weight of earlier thought. On this reading Descartes is seen to be still a prisoner of Aristotle for whom biological nature above all and even nature in general cannot be adequately understood except in the light of the good.

There is at least one other sentence in Descartes' account that might be brought to bear in support of the hypothesis that he is guilty of an inconsistent mechanism. He says that *le vrai usage de la respiration*—

rendered somewhat too freely as "the true purpose of respiration" in Paul Olscamp's recent translation—"is to carry sufficient (*sic*) fresh air into the lungs to cause the blood which . . . has been . . . transformed into vapors to thicken and become anew converted into blood . . . , without which process it would not be fit to serve (*propre à servir*) as fuel for the fire" in the heart. The word 'purpose' being for us a red flag signalling the presence of teleological as opposed to mechanistic considerations, it is doubtless too strong in the present context, though there is enough 'indeterminacy of translation' to prompt the suggestion that the more neutral word *usage* may merely obfuscate the teleological import of the passage. The last clause of the sentence seems designed to explain how we can identify the true use (usefulness?) of respiration through an appeal to what Aristotle calls 'hypothetical necessity'. If there is to be an (healthy) animal the fire in the heart must be fed but if that is to occur blood must be sent to the heart and if that in turn is to occur the respiratory apparatus must send enough fresh air into the lungs. So we thereby come to know the (final) cause of respiration: "respiration" takes place in order that enough fresh air be sent to the lungs in order that . . . , almost uncannily recalling Aristotle's *Parts of Animals* 642a31-642b5. But one would suppose that final causes are precisely the sort of pseudo-explanatory devices that Cartesian mechanism is calculated to eliminate.

II

Although the tone of my exegetical remarks has been largely negative, it has not been my intention to criticize Descartes. Quite the contrary. That normative if not teleological considerations are required for an adequate understanding of biological functions will prove to be my own thesis, but I do not suppose that Descartes' own ambiguous example can provide more than a first clue toward accommodating that thesis within a mechanistic framework. If it were merely a matter of 'exonerating' Descartes from the charge of permitting his mechanism to be infected with the residual traces of an antiquated teleology, the readiest defense would be to insist on the highly informal, even rhetorical, character of the *Discourse on Method,* in which there are actually very few pages of straight philosophy, let alone straight science, compared with the bulk of the work, which is filled out with an extended, programmatic manifesto. It is enough for Descartes' purposes to convey the flavor of a strictly mechanistic account without being overly fastidious at every point, seeing that only with the greatest difficulty can one prevent an informal account

of biological matters from lapsing occasionally into at least a quasi-teleological idiom, value-judgments never being far to seek.

Quine's own practice is instructive. After expelling value judgments from scientific theory he does not hesitate to readmit them in his own informal, quasi-biological account of the growing infant as budding scientist in swaddling clothes.

> We are working up our science from infancy onward. Each of the leaps of language learning that I have pictured is a private little scientific revolution . . . If the leap is one that conduces to simplicity in the child's evolving conceptual scheme, then normatively speaking it is good scientific method on his part, however unconscious. If it is a short leap, then again it is good, on the score of conservatism.[2]

In reply to the objection that Quine is here engaged in philosophy not science one has only to remember that 'total science' for Quine refuses to honor the distinction on the highest level, and if Quine's 'psychology' of infant learning is correct I fail to see why the normative considerations he adduces should not be accepted as belonging to science proper.

The distinction between a good argument and a bad argument, between good evidence in support of a theory as opposed to the contrary: these are matters in which scientist and philosopher alike have a vested interest, and they are thus committed to holding that there are objective criteria for evaluating all three—arguments, evidence, theories. One may indeed ask with Hume whether we have good grounds for believing that the sun will rise tomorrow but if one answers with Wittgenstein, "A good ground is one that looks *like this*" (*Phil. Inves.*, I, 483), he cannot but allow that at least some value judgments are objectively true. Again, it will not satisfy a Quinean holist if we accept value judgments on a metascientific but not a scientific level, the distinction between the two being 'sublated' in total science. And if it is agreed that "the general question of scientific method" is "how best to develop an inclusive scientific theory" and that "guiding considerations are simplicity and conservatism" (*Op. cit.*, p. 137) we have then objective criteria in hand for evaluating "the private little scientific revolutions" of the growing child (assuming indeed that Quine is right in postulating their existence). That Quine as a pragmatist should be seen undertaking to bridge the fact-value gap is what one would expect, and one can only be surprised by his choosing to stand officially "aloof" from value judgments. In his brief but thematic discussion entitled "Values" it turns out that it is really only aesthetic and moral values in regard to which he opts for non-cognitivism, where by "moral values" he means "the moral law" and "the 'good' of

righteousness". But there is a lot of objective, natural good left over, as when he speaks of the "benefits" that "language confers . . . through the conformity of speech habits", not to mention how a man "may stand to gain much from wicked behavior once the sanctions are dropped."

It must be admitted that Quine's armchair speculations concerning child development may be felt to be much too whimsical to afford any solid basis for supposing that hard science can be normative. Modern physics since Galileo being conceded to be 'value-free', the test case must be the science of biology proper, and here one might well feel Darwin to be the decisive figure in finally discrediting the whole teleologico-normative framework of Aristotelian thinking, traces of which we found in Quine as well as Descartes. Startling indeed is the sharp contrast between the dry language of Descartes and the luxuriance of Darwin's on these very matters. Darwin characteristically expresses himself as follows.

> . . . the more complex organs and instincts have been perfected . . . by the accumulation of innumerable slight variations, each good for the individual possesser . . . the instinct of each species is good for itself but has never, as far as we can judge, been produced for the exclusive good of another . . . Natural selection . . . leads to the improvement of each creature . . . What limit can be set to this power . . . favoring the good and rejecting the bad? . . . as pollen is formed for the sole purpose of fertilization . . .[3]

If modern natural science eschews value judgments Darwin certainly was never informed of that fact. Nor can his profligate use of the normative idiom be entirely dismissed as mere literary turns of phrase, though doubtless some of it may be thereby explained away. That there is an objective, natural good for animals and even plants produced by the Invisible Hand of natural selection, Darwin never doubts; and his example should be sufficient to undermine the fashionable thesis that modern natural science as such, even when left to its own devices unmonitored by specifically philosophic scruples, has no use for the distinction between good and bad.

Lest it be felt that Darwin's example is not quite up-to-date enough the following is a typical expression of current thinking that the working biologist is prepared to find platitudinously congenial: "To understand behavior we must look first at certain aspects of the nervous system . . . in recording changes in the environment, in assessing this information and in coordinating the various muscular activities that will best serve the economical operation of the animal"[4], where it is being tacitly assumed that we have criteria in hand for recognizing occasions when, owing perhaps

to some malfunctioning of the nervous system, there is a failure in muscular coordination and the activities of the animal no longer "best serve" its interests. Here it strikes me that we are very much on the right track and that there is no need to apologize for the normative element. Granted that Darwin's own language is too undisciplined to inspire much confidence, the picture of natural selection as "favoring the good and rejecting the bad" being conceptually as well as morally objectionable, the fact that some sort of retrenchment in the use of the normative idiom is certainly required need not lead one in panic to jettison it entirely.

Still, there are deep conceptual misgivings that need to be allayed of which two are perhaps salient. First, there is the following, now almost classic argument that must be confronted. (1) Natural science can be normative only if 'good' can be defined in naturalistic terms, (2) But G. E. Moore has shown that any such definition of 'good' will fall foul of the naturalistic fallacy. Therefore (3) natural science must shun all value judgments. Although I accept the second premise as true I reject the argument on the ground that the first premise is false. How natural science, biology in particular, can be a normative discipline even though 'good' resists being naturalistically defined, must then be clearly explained.

A second source of misgivings may be summed up as follows. Although Darwin's example should suffice to show that a reputable scientific theory, and even a scientific theory characteristic of modern as opposed to ancient science, may well be framed in terms of normative considerations, one may be prepared to allow value judgments to figure only in the less rigorous, more informal versions of a scientific theory even while one ruthlessly expunges them from the more austere. There being various "grades of austerity" available for the formulation of science, one may choose to take one's lead from Quine who is willing to acknowledge the existence of numbers (i.e. to quantify over them) on any level of discourse with the exception of the highest, most authoritative, where he eliminates them, replacing them by sets. Normative and numerical considerations might thus be taken as being roughly on a par, both being allowed as quasi-scientific. In an eirenic spirit, then, let it be granted that the normative and teleologico-normative idioms are suitable enough in the less austere formulations of the life sciences (one is thus allowed to say that milk is good for cats), and let us only ask whether they are to be retained in the stricter formulations. But serious doubt may here arise to whether the life sciences can even in principle admit of any highly austere formulation without at once being submerged by physics and chemistry.

If the natural sciences are finally reducible to physics one is entitled to expect the normative to be swallowed up along with biology itself.

This second source of misgivings being physicalistic while the first was 'analytical' in import, it is the combination of the two rather than either taken separately that is felt to be peculiarly coercive in contemporary philosophy. I shall deal with both.

III

"Early versions of mechanism," writes David Hull in his discussion of "the most primitive formulations" of it prompted by Descartes, "were characterized by two basic tenets: the belief that all of science could be derived from the science of mechanics and the resulting conclusion that living creatures could be treated as machines pure and simple."[5] If the first of these two tenets is expressed by Descartes when he says the rules of mechanics are the same as those of nature, the second is launched with his doctrine of the *bête machine*, and indeed after completing the physiology of *Discourse* V. he asks us to compare the workings of the human body to those of various, ingenious automata made by men and thus to "consider this body as a machine . . . made by the hands of God." However much it may appear as a mere "resulting conclusion" derived from his first tenet, it is precisely the novelty of his second that will be found to help explain what we could only take to be 'lapses' from the rigor of his avowed mechanism. What is involved here for an animal to be "treated" or "considered" as a machine when literally speaking it is actually nothing of the sort? Allowing for stricter and looser uses of the word, 'machine' in its primary signification is true only of (some) artifacts, and every animal of our acquaintance (if only as a matter of fact) fails of being such. Which is not to deny that powerful illumination might ensue upon our regarding an animal *as* a machine.

With wheels, pulleys, pistons and meshing gears a standard machine might well be thought, almost by definition, to lend itself in an entirely transparent manner to a purely mechanistic account, and as such it might be felt to serve admirably as a model for understanding living organisms which otherwise might be taken (again by definition) to require some sort of 'vitalistic' explanation. But why should the machine analogy be felt to be peculiarly appropriate in the case of animals when no one would suggest that a rock or an icicle might be "treated" as a machine despite the fact that they are much more obviously open to a mechanistic account

than even the simplest animals? Paradoxically enough, it is precisely because a machine cannot be understood in purely mechanistic terms that it recommends itself as a model for the understanding of animals. To understand a machine one must understand its function, its purpose, what it has been designed to do; and in the light of that understanding one can determine whether the machine is working well or working badly, whether perhaps it is not working at all (though wheels may be spinning furiously), whether having broken down it can be restored to good working condition and so forth. Built into the very concept of a machine, what sort of machine it is, are the criteria for evaluating both its external performance and its internal operations. Far from being alien to the concept of a machine, normative considerations are found to be linked to it logically by the most intimate connection.

The whole Aristotelian schedule of the four causes—final and formal quite as much as efficient and material—applies to a machine in a perfectly straightforward, philosophically unproblematic fashion. It is only when one undertakes to understand nature on the model of art, as Aristotle explicitly and Descartes implicitly is seen to do, that serious doubts cannot but arise. No radical break with Aristotelian thinking, the *bête machine* and *homme machine* hypotheses can only be understood as efforts (never brought to a point of full clarity) to reinstate the teleologico-normative standpoint within a mechanistic world view. No wonder, then, that considerations of good and bad should surface in Descartes' mechanistic, i.e. machine-like, account of the biological functioning of the heart, and it is only surprising that they should be as fugitive as they are. Even in the much more extended *Treatise on Man* (107 pages in Clerselier's first French edition) of which the physiology of *Discourse* V. is merely an abridgment Descartes is seen to be remarkably abstemious in his use of the normative idiom and not merely if his practice is compared with the prodigality of Darwin but even apart from any such comparison. There may in fact be only two explicit value judgments that creep into the whole work, one where he tells us what happens "when the liver is *bien disposé* and elaborates perfectly the blood that goes to the heart", the other nearby where he tells us what happens when "the spleen which is *destinée* to purge the blood of parts least suited (*propres*) to be enkindled in the heart is *mal disposée*" (Adam and Tannery edition, Vol. 11, p. 168).

Much easier to find, however, are implicit value judgments embedded in what can only be described as at least quasi-teleological contexts, as when we learn that "the changes that occur in the size of the pupil serve to moderate the strength of the vision; for it (the pupil) needs to be

smaller when the light is too bright, in order that so many rays do not enter the eye that the nerve is damaged thereby." (*Ibid.* p. 156). The concept of damage in animal as well as machine presupposes a standard of good and bad, for one can recognize a nerve as being damaged only after its function has been identified and one can contrast its present state with some other which constitutes its being in good functioning condition. How the concept of damage actually guides ongoing biological research can be seen in recent efforts to prove that the thymus is an endocrine gland. As a general principle of methodology, Allan Goldstein and Abraham White observe that "one postulate that must be fulfilled experimentally for establishing an endocrine role for a given anatomical structure is evidence that physiological and biochemical deficencies ensue following the extirpation of the organ or tissue in question." Such "deficiencies" have been sought by investigators in studies of experimental thymectomy for more than a hundred years but only lately has it been verified that "deleterious results . . . ensue as a consequnce of removal of the thymus,"[6] and in this way the thymus has been found to have an endocrine function.

If the concept of function with its built-in normative import is indispensable for understanding the inner and outer operations of both animal and machine, the logic of "functional words" according to Richard Hare is infected with certain "peculiarities" that are equally worth noticing. From Hare's noncognitivist standpoint they must be very peculiar indeed if they can give rise to the one exception to his otherwise adamantly held principle that evaluative statements cannot be logically derived from purely factual ones. Conceding that "we can indeed say that if an auger will not bore holes at all then it is certainly a bad one", he recognizes that in the present instance "we are handed on a plate, in virtue of the meanings of the words used, one of the necessary criteria of a good auger."[7] But it is only in this very marginal sort of case that Hare (in effect) allows that the fact-value dichotomy can be cognitively overcome, and indeed in a later discussion where he reaffirms the absolute impassibility of the fact-value gap he is forced to conclude that "the word 'hygrometer' is . . . not purely descriptive", on the ground that "to know the meaning of 'hygrometer' . . . we . . . have to know something about what would justify us in commending or condemning something as a hygrometer."[8] It then follows on this account that the apparently factual sentence 'There are three hygrometers in the storeroom' can never be objectively true, but this is clearly unacceptable if for no other reason than that we are left with no entity on hand, no 'something', with its own independent identifying essence (i.e. persistence conditions over time),

which can be picked out in the first place and then evaluated *as* a hygrometer. Once the primary sortal 'man' has been used to pick out an entity it can be evaluated as a soldier by means of the secondary sortal 'soldier'. In Hare's case there is no such primary identifying sortal. 'Instrument' will not do, seeing that it is a generic term which can be no more 'purely descriptive' than the species which fall under it (e.g., 'auger', 'hygrometer').

That 'functional words' with their 'peculiarities' should be a source of embarrassment for the noncognitivist might be easily taken to be of only marginal interest, as bearing merely on considerations of instrumental (as opposed to intrinsic) value. By no means marginal however in biology, functional words play a central role in the life sciences, though one must not expect the criteria of evaluation with which they are laden to be handed to us quite on a plate. In the lexicon of Goldstein and White 'endocrine gland' appears to be a functional term whereas 'thymus' clearly is not. An 'anatomical structure' that can be identified independent of function, the thymus can be shown to be an endocrine gland only if (the right sort of) physiological and biochemical deficiencies result from its removal, implying that 'endocrine gland' is a functional term with a built-in normative force. Moreover, "it is presently indicated," according to Goldstein and White," that not only do thymic-adenohypophyseal endocrine secretions alter lymphoid tissue structure and function but balanced interrelationships among these hormonal factors may comprise homeostatic mechanisms for modulating the functions of the lymphoid system" (*Op. cit.*, pp. 467-68), which I take to imply that one function of the thymus may well be to support certain homeostatic mechanisms. Whether applied to biological or merely artificial processes, 'homeostatic mechanism' is certainly another functional term: one can always ask of any such mechanism whether it is working well or not, and the criteria for answering the question are always available at least in principle.

'Machine' having been seen to be a portmanteau word packing together mechanism and teleology, the term 'mechanism' itself is richly ambiguous, having a normative as well as its non-normative use. Strictly non-normative when it is applied to the classic case of billiard balls colliding and rebounding in accordance with the laws of Newtonian (or Cartesian) mechanics where the concerns of the billiard players are waived, the term has its normative application when the watchsmith remarks on various mechanisms inside a watch, these being or failing to be in good working order. Let us then distinguish between mechanism$_1$ in its non-normative and mechanism$_2$ in its normative sense, and there is then no paradox if we say that machines cannot be understood in purely

mechanistic$_1$ terms, though they can of course be understood in mechanistic$_2$ terms. That the *bête machine* hypothesis is at least prima facie vulnerable to the charge of committing the fallacy of equivocation, conflating the two senses of the word 'mechanism' in pretending to explain animals purely mechanistically$_1$ while in fact it regards them mechanistically$_2$, may well be especially evident today when the hypothesis can be said for the first time to have really come into its own as a regulative principle in such accounts as the following.

> The Respiratory Control System. A step input to the respiratory system of carbon dioxide in inhaled air causes the arterial and brain tissue CO_2 to rise. In response, both cerebral blood flow and alveolar ventilation increase in an effort (*sic*) to maintain the brain CO_2 concentration at the normal level . . . it is evident that the respiratory system regulates CO_2 and O_2 levels through negative feedback.[9]

Explicitly assimilated to engineering models, the cybernetic story here of one homeostatic mechanism$_2$ as a goal-directed control system might readily incline a cynic to feel that it is precisely by masquerading under cover of an aggressive, hardheaded mechanism, i.e., mechanism$_1$, that the teleologico-normative standpoint of Aristotle is seen to be creeping back into favor, no secret being made of the fact that the object is

> to describe how living plants and animals work in the same sort of way that we describe how non-living machines . . . work . . . adjusted and controlled in response to changes in the external environment . . . for the *well-being* of the whole animal or plant . . . From this point of view regulation and control in the living world depend on the existence of automatic control systems. Engineers have made extensive studies of the general principles on which they must be designed if they are to work properly.[10] (my italics)

The merest small talk of the natural sciences, such *obiter dicta* might, however, be felt to be 'privileged' utterances that are not to be looked at too closely, being not much more perhaps than incidental chit-chat ushering in the applied mathematics comprising the substantive material of the book.

The mention of mathematics suggests indeed the following parallel. Mathematicians having proved the existence of a prime number greater than 17 but less than 29, is it some material object or physical entity that they have proved to exist? Virtually forced to answer in the negative, most mathematicians will feel uneasy if it is insisted that what has been proved to exist by their science must then be something immaterial or non-physical. Nor is it at all clear how their existential idiom is to be purged of its ostensible reference to such abstract entities. If the natural

language of the mathematician is thus seen to be Platonistic, it is no less the case that the natural idiom of the biologist is Aristotelian, and the power of that idiom is not to be exorcized simply by affecting a tough-minded materialism and saying with Claude Bernard that "the organism is merely a living machine." Merely! The parallel can be carried a step further. The fact (as Quine has stressed) that no one has succeeded in providing anything like a convincing nominalistic reduction of mathematical discourse encourages us to argue that the normative idiom in its turn is not to be dispensed with in the life sciences. Basically, it is not to be dispensed with because the concept of function (and malfunction) is indispensable, and once the respiratory system (say) has been recognized to be a homeostatic mechanism, the function of which is that it "regulates CO_2 and O_2 levels through negative feedback," one is already committed to objective criteria for evaluating its performance. Just as an auger that will not bore holes at all must be a bad auger, so too, a respiratory system that fails to regulate CO_2 and O_2 levels must be said to be in bad condition.

IV

Suppose that at t_1 you notice a cat lying on the mat alert and intense but a Martian physicist—lacking perhaps the very concept of an animal if such extreme conceptual deprivation can be imagined—is only aware of a pile of (interlocking) atoms in that region. A few hours later at t_2 the physicist reports that the same pile of atoms, indeed the same atoms, has persisted *in situ* from t_1 to t_2, though it is now perfectly plain to you that there is no longer a cat on the mat, not even a dead one. In the interim the atoms constituting the cat have become so radically re-arranged that no form of life or death is to be discerned in the area. Although the physicist may well be prepared to trace the career of each atom throughout the interval from t_1 to t_2, he has failed to recognize that a certain entity, namely the cat, has ceased to exist, never having been aware of it in the first place. Failing to see the wood for the trees, his observations have been too 'atomistic' (in both the conceptual and physicalistic senses of the word) when what is required for the recognition of an animal is a holistic kind of apprehension that will grasp a certain kind of pattern, an *eidos*, enabling us (if the pun may be allowed) to eidentify those entities (they are not piles of atoms) of which the predicate 'is a cat' or, more generally, 'is an animal' is true. Only after such entities have been eidentified can one go

on to pick out those items of which the (observational?, theoretical?) predicate 'is a respiratory system' is true.

That life can only be understood holistically is a familiar enough refrain in modern thought at least since the time of Goethe, but as the mention of the poet's name suggests, it has almost always tended to be associated with some sort of romantic irrationalism in its protest against mere 'analytical', i.e. atomistic, modes of thought ('we murder to dissect'), and we are certainly not encouraged to make much headway in unpacking the concept of life. Built into it in fact is a whole battery of sub-concepts of which those of function and malfunction are only two and not the most immediate or most 'phenomenal' from the standpoint of recognition. If an animal is to be identified in the world at large we have only bodies in motion to go by and, granted that these have been picked out, it is the concept of action (as opposed to mere movement) that must be brought to bear upon them. Paraphrasing Wittgenstein's question, we have to ask what is the difference between a cat raising its paw (always an action) and a cat's paw rising (sometimes a mere movement). I take it then that animals are recognized primarily by their actions (even in the case of amoebas it can be said that "virtually all stimuli that are not food elicit withdrawal reactions")[11] and that actions in turn are recognized by means of an action *eidos* or action pattern, though if we wish to know further how to distinguish actions from mere movements we are urged by Wittgenstein to rest satisfied with the following 'explanation'. To acquire the predicate 'is a cat raising its paw', that is, simply to understand what those words mean, entails being able to use it correctly in different grammatical constructions in accordance with the conventions of the language, but this intra-linguistic competence is not enough, there is also the extra-linguistic requirement of being able to use the predicate correctly to pick out those entities of which it is true as well as those of which it is false. Precisely how one is to go about doing that, what criteria one is to employ in exercising this second competence, one simply learns *pari passu* in learning the meanings of certain words (i.e. mastering certain conventions): it is conceptually impossible to learn the latter (the meaning of the words 'action' and 'movement) without also learning the former (how to distinguish actions from mere movements). So it all comes down in the end to convention, and there is no mystery.

But there is a mystery. If we have in hand criteria for distinguishing actions from mere movements one would naturally suppose that such criteria (like all criteria) might in principle at least be brought out into the open and explicitly formulated, but in recent action theory it is invariably

assumed that "this schematism," if I may press into service Kant's idiom here, "is an art concealed in the depths of the human soul, whose real modes of activity nature is hardly likely ever to allow us to discover, and to have open to our gaze." 'Schematism' is the right word, for on being confronted with the movements of bodies we may be said to schematize them in such a way as to eidentify some of them as constituting actions and the rest as movements merely. The action *eidos* being then the schema of life it must be asked whether that schema is represented in the human nervous system by some underlying mechanism$_2$, in the normative sense of the word, seeing that it is always liable to malfunction, and whether it can not then be fully described. Above all, we need to know if there is a formula which reads roughly as follows: "If and only if bodies are moving in such and such ways . . . are there animals engaged in actions", with the blank being filled with straight kinematics. Although the discovery of that formula would be of the first importance it would doubtless prove to be so complex, filling thousands if not millions of pages, that nothing might count as having it "open to our gaze." If the life formula should fail to accommodate protozoa, no matter. As the name applies, protozoa may well be only proto-animals that only engage in proto-actions or even merely proto-engage in actions. One may then expect the concept of action to apply to them with highly diminished force, and the life formula may recognize them only as degenerate cases of animals (in the mathematical sense of the word).

If the life formula may be said to perform a reduction of actions to movements and thereby of biology to physics, it is only in the sense that it specifies precisely what bodily movements (some occurrent, others subjunctively counterfactual) constitute actions. Notice, however, that the concept of action, indeed the word 'action' (as contrasted with 'mere movement') remains irreducibly present in the formula, on its right side: these items (movements) constitute those (actions). The case is not at all comparable to that of Quine's eliminating numbers and replacing them by sets (numbers being neither constituted by nor identical with sets). Reduction, then, by elimination is not to be confused with reduction by specification. I can readily imagine someone, our Martian physicist perhaps, fully mastering the life formula on the mathematico-physical level without having any idea that it is the formula of *life*. He would learn the formula in the mutilated version, "These movements . . . comprise the Q class of movements" where 'Q' is a meaningless proper name. The life formula is only adequately understood when it is seen to bring into line both the concept of life and its kinematics.

If conceptually speaking an animal is recognized primarily by its

behavior and animal behavior is characterized above all by the performance of actions, the central core of animal actions must be understood in the light of the concept of purpose. Not that every time a cat raises its paw some sort of purpose or even quasi-purpose need be read into the act. The cat may be raising its paw for no reason whatever (as we say), mere idle activity. But such aimless actions (they are not, like the kneejerk reflex, mere movements) can never comprise the central core of animal behavior. Purposive behavior then is salient, and corresponding to the concept of purpose in enabling us to schematize the external movements of an animal is to be found the concept of function in allowing us to schematize those that are largely internal to it. The cat is in motion, it is pursuing the mouse, using both its eyes and its feet, the former for the purpose of determining where the mouse is, the latter for the purpose of reaching it. The eyes and feet are then identified as organs, each being indeed an *organon*, i.e. an instrument, the eyes having a sensory and the feet an executive function. The organism itself is thus seen to be an instrumentalized entity, the external organs being sensory and executive instruments and the internal organs (heart, liver etc.) providing a support system.

It remains true, however, that if one should wish to introduce a child to the word 'instrument' or 'tool' one would not adduce eyes, feet, teeth or claws as paradigm cases but rather an axe, a hammer or an auger, artifacts. It is only by analogy then that the organs of an animal can be said to be instruments; and the term 'function' itself must first be learned as embedded in such expressions as 'the function of a hammer', 'the function of an axe', only later being used in connection with the heart or liver. When Aristotle undertakes to explain what soul is he invites us to imagine a natural, i.e. non-artificial, axe which is presumably alive (*De Anima*, 412b13) and perhaps to be pictured as running through the forest chopping away on its own, very much a *bête machine*. As the function of an (artificial) axe is to chop or cut, so, says Aristotle, is chopping or the capacity to chop the soul of the living axe or axe-animal, and by means of this analogy we are brought to see that soul is simply natural function, the nutritive soul of an animal being merely the nutritive function, the sensory soul the sensory function and so on. It is scarcely a major step now to speak of various mechanisms$_2$ operative in the organism, digestive, respiratory, perceptual etc., where again the term 'mechanism' even in its normative import is derived by analogy from such artificial mechanisms$_2$ as those that are to be found in a watch. That analogical thinking should play such a decisive role in the life sciences need not cause alarm. The mathematical concept of a set is derived by analogy from our ordinary use

of 'collection', 'set,' 'heap,' etc., and it has been suggested that the physical concept of a force is related in the same way to our ordinary use of the word 'force'.

Mere unaided sense-perception will not suffice to enable one to identify the digestive mechanism of a crocodile. A richly schematized mode of attention is required, and built into the criteria for recognizing such a mechanism in the first place are to found the further criteria for evaluating its performance. Only by understanding the function of the digestive mechanism can one so much as pick it out in some organism, and on the basis of that understanding one is now in a position to investigate whether or not it is in a good, healthy condition. Committed to the study of such mechanisms, the physiological sciences come to sight as invincibly normative, above all that particular physiological science known as pathology, which investigates the various diseases with which animals (and plants) are afflicted, always presupposing in its inquiries an objective standard of health. In a commensensical sort of way no one doubts that an animal in a diseased condition is in a bad condition and that an animal in a healthy condition is in a good one, but these merest truisms—life tautologies, as we may term them—have been felt to be extraordinarily difficult to absorb into the world view of modern science, in part perhaps becuse they have seemed to be most at home in the framework of an Aristotelian 'naturalism' against which modern science with its emphasis on mechanism$_1$ resolutely set its face from its inception. By replacing Aristotelian 'naturalism' by Kantian schematism I have tried to show how the life tautologies can be accommodated in a world of bodies whose kinematics is indeed to be understood primarily in terms of mechanism$_1$ but secondarily, and in a limited sector, in terms of mechanism$_2$ where a certain sort of entity, namely an organism, is to be recognized and identified by means of the concept of life and its associated schema, the latter perhaps capable of generating the life formula, the former on being unpacked certainly yielding the contrast between mere movement and action, purpose and function, health and disease.

Because of its background in Aristotelian 'naturalism' it might easily be thought that our functionalist theory of value might well be guilty of some version of the naturalistic fallacy. If the digestive mechanism of a crocodile is fully functional and thus free of all forms of malfunctioning it is doubtless in good condition, and it might then seem that we are in effect defining the good as the functional. Although that slogan is acceptable enough in an informal sort of way it would be wiser to define 'good', if define it we must, following John Searle's proposal as "meets the criteria or standards of assessment or evaluation".[12] The definition is certainly

not naturalistic, and there is no room for raising Moore's 'open question'. One cannot protest, "I agree that Newton's theory fully meets the (correct, objective) criteria for evaluating theories, but is it after all a good theory?" The example was not chosen at random, having been fetched fairly close to home. Engaged as we are in the theoretical enterprise of philosophy and science, we are pragmatically committed by our very endeavor to there being objective criteria for the evaluation of theories, arguments, and evidence, and when we insist that one theory is better than another it could only be self-stultifying to concede that our assertion lacks a truth-value, merely because it is a value judgment. That the theoretical enterprise as such is pragmatically committed to the acceptance of certain principles which otherwise (i.e., when viewed in the mode of pure objectivity) might be felt to be contestable, will be remembered as one of Plato's favorite themes (cf. *Parmenides* 135 b-c, *Theaetetus* 170-e-171, 183 a-c, *Sophist* 249b-c), rooted finally in the Socratic imperative of self-knowledge; and in arguing here that we are pragmatically committed to there being objective criteria of evaluation in philosophy and science I am to be seen as merely practicing a form of Platonic pragmatism.

We are not, however, committed to the meaning of 'good' being definable, recent philosophy having indeed shifted attention away from questions of meaning to those of truth-conditions. Thus it is no longer the meaning of 'know' that is studied but rather the truth-conditions of someone's knowing that p. So, too, we must ask after the truth-conditions of something's being a good F. Well, if there are objective criteria for evaluating F's (say C_1, C_2 and C_3) and if this F fully meets those criteria, it logically follows that this F is a good F. One has merely to throw in the necessary hence otiose premise that if some F fully meets the objective criteria for evaluating F's it is a good F, in order to provide a recipe for producing a formally valid argument that overcomes the fact-value hiatus. The recipe being merely emptily formal fails to give any indication of how it is to be applied with substantive content, but for that one must look elsewhere—above all to the life sciences, especially pathology, where objective criteria of evaluation are readily available, packed into the concept of function, and where the descriptive and normative idioms are woven together. Simply to describe how the respiratory system of Dobbin is functioning (or malfunctioning) is impossible unless it is recognized to be a mechanism$_2$ which may be working well or working badly. How loaded is even so simple a statement as that Dobbin is breathing with difficulty! One may indeed decline to recognize Dobbin in the first place, preferring merely to investigate the pure kinematics (and dynamics) of micro-bodies. But Dobbin is there to be recognized, by the eidetics and

schematics afforded by the concept of life, and once one does pick out an animal from amid the welter of the narrowly physicalistic, a whole conceptual scheme, not least of which is the normative component, is found to be activated.

One advantage of the present approach is that it contains an escape clause which allows us to adopt a quasi-noncognitivist position in regard to a wide range of value judgments. Suppose (regressing for a moment) that in asserting some joke to be a good joke one is to be taken to *mean* that the joke fully meets the objective criteria for evaluating jokes. Suppose further that there are no such objective criteria. Must the value judgment then be said to be false? Not if we adopt a Strawsonian reading which takes us to be merely presupposing, not asserting, that objective criteria obtain. The case then will be on all fours with saying that the present king of France is bald when there is none. On a Strawsonian (though not a Russellian) reading the value judgment proves to be neither true nor false. In any event, in saying of some F that it fully meets the objective criteria for evaluating F's one is certainly commending it (as an F), and we can thus honor Hare's insistence on the commendatory force of 'good' without depriving value judgments of their factual import.

In his sixth *Meditation* Descartes in commenting on the dropsical patient whose craving for water is badly out of line with his bodily needs, as we at least have no hesitation in saying, seeing that need and desire are no longer functionally integrated, observes that the laws of nature, i.e. the rules of mechanics, hold sway quite as much in the unhealthy, diseased man as in the healthy one, and therefore (he infers) it is only by "comparing" the former in purely "extrinsic" fashion "depending on my thought" to a badly constructed clock and the latter on one well designed that it may be said—in Aristotelian style and altogether illegitimately—that nature has somehow gone wrong, failed in its purpose, or been "corrupted" in the former but not in the latter case. I hope to have shown how these objectionable turns of phrase characteristic indeed of Aristotle can be effectively eschewed without, however, going so far as to deny (as Descartes is on the brink of doing) that the difference between health and disease is fully objective and in consequence open to scientific research. It remains true, however, that the normative distinction between health and disease cannot be adequately explained by the rules of mechanics or understood in their terms, and if we had no other principles to fall back on we should be forced to dismiss it from the purview of sober science (always bearing in mind that the H class of physical configurations is objectively distinguishable from the D class, where 'H' and 'D' are mere brute proper names to the Martian physicist, while they signify to us

'health constituting' and 'disease constituting' respectively). But though it cannot be explained in terms of mechanism$_1$, it can indeed be understood in terms of mechanism$_2$, and Descartes' own hypothesis of the *bête machine*, itself modeled on artifacts, has shown us the way. It is ironical to find that in his mature thinking Descartes failed to take it seriously enough, having become dimly aware in the sixth *Meditation* that "to consider the body of a man as being a sort of machine", far from involving a radical break with Aristotle, opens the way to "extrinsic" normative considerations which his mechanism is ill equipped to handle. It may be doubted, however, whether Descartes ever came to any clear understanding with himself on the issue. Four years before his death when he has occasion to rehearse his mechanistic account of the heart and its movements in the *Passions of the Soul* scarcely any trace is to be found of the teleologico-normative idiom which intermittently infects *Discourse* V. and the *Treatise on Man*, but somewhat incongruously he continues to speak of "the bodily machine" and even to "judge that the body of a living man differs from that of a dead man just as does a watch or other automaton" which is fully operational "from the same watch or other machine when it is broken and when the principle of its movement ceases to act" (*Passions*, I, articles 6 and 7).

The case of the dropsical patient in his distempered state proves to be for us a positive source of instruction. It is precisely in the functional integration of an animal—wolf or lamb, lion or fox, serpent or dove—that what it is for it to be in good, healthy condition will be found to consist, each in its own species-specific way. But if the (intrinsic) good of an animal is to be seen above all in its functional integration, it is no less the case that the animal is constituted as an entity in the first place by an even 'prior' 'act' of integration, as Sherrington brings out with peculiar vividness when he asks how it is that "the separate units of an animal body are welded into a single whole and from a mere collection of organs there is constructed an individual animal,"[13] inviting us somehow to imagine the organs of an animal as being initially (in some logical though not temporal order) laid out before us as mere *disjecta membra* that await their appropriate consolidation. Where Aristotle's answer to the question is that it is thanks to soul that such welding is found to obtain, Sherrington's is that it is owing to the "integrative action of the nervous system". Both answers are less than satisfactory, Aristotle's because it is scarcely more than a place-holder for a content-filled account, Sherrington's because it begs the antecedent question as to how the separate units of the nervous system are themselves welded into a single whole. Moreover, according to Ragnar Granit, "his 'integrative action' . . . is the knitting together of

cells activated across synapses for a common goal",[14] which is rather too teleological to suit us in quite that form, though it can be rescued if it is understood in the light of *our* "goal" being literally to effect "the knitting together of cells" with the express aim of having an animal "constructed" by us out of a "mere collection of organs."

Let us imagine a laboratory well stocked with a varied assortment of mature hearts, lungs, claws, eyes, feathers, etc. or, in somewhat more abstract terms, let there be here a shelf of digestive mechanisms, there a supply of sensory mechanisms, over there a bank of muscular mechanisms, etc., not to mention a wide variety of nervous systems—all in a state of high operational readiness, and judiciously picking and choosing among them let us construct some sort of animal—never perhaps seen before, being as strange as any the poets might feign or the twisted imagination of an H. P. Lovecraft relish—by organizing organs and welding physiological mechanisms together, where the very word 'weld' is redolent of the machine shop. It is not, however, the Baconian program of a technology of life (in the *New Atlantis* it is said that "we make a number of kinds of serpents, worms, flies, fishes, of putrefaction") that particularly interests us nor even the purely theoretical discovery of some mechanical or paramechanical vital adhesive that binds the components of an animal together. Our concern is less physical and more logical, conceptual, eidetic, schematic, mathematical, the aim being finally to express in some master formula the principle of functional integration that constitutes an animal in the first place. The master formula would state in entirely general terms how bodies (molecules even more than organs) must be arranged if an animal (be it healthy or unhealthy, familiar or exotic), is to result. There will then be a second formula which will describe what must obtain if there is to be an animal in good, healthy condition, and the two formulas will be related in important ways. Here again analogy is scarcely to be avoided. There is (I presume) one formula that describes what is required if there is to be any old auger or hygrometer and another formula that specifies what is needed if there is to be a good auger or a good hygrometer. In the case of animal as well as machine there will then be both a primary and secondary principle of functional integration, where the latter somehow delivers 'more' of what the former provides, the former merely laying down the conditions of life, the latter the conditions of the good life.

According to Judith Swazey, Sherrington wished to "explain the functional unity of motor behavior" (*Op. cit.*, p. 168), but any such explanation must have a conceptual, indeed a normative as well as a physical side to it. What counts logically as motor output in an animal,

action above all, is conceptually as well as physically bound up with sensory input and their functional integration by the nervous system: action to be action must be guided by sensory information, and sensory input if it is to be *sensory* input must be (at least potentially) manifested in behavior where "behavior as behavior", in Tolman's words, "reeks of purpose and of cognition."[15] But that integration may be more or less effective, the sensory and motor mechanisms may be well coordinated or badly out of line with one another owing to some malfunctioning of the nervous system, and it is thus seen how the normative and descriptive idioms are themselves functionally integrated when it comes to describing, and evaluating, how the organs of an organism are organized.

Syracuse University

NOTES

[1] W. V. Quine, *The Roots of Reference* (La Salle: 1973), p. 49.
[2] *Op. cit.* p. 138
[3] Charles Darwin, *The Origin of Species* (Modern Library edition), pp. 353, 186, 98, 360.
[4] V. G. Dethier and Eliot Stellar, *Animal Behavior: Its Evolutionary and Neurological Basis* (Englewood Cliffs: 1964), p. 3.
[5] David Hull, *Philosophy of Biological Science* (Englewood Cliffs: 1974), pp. 129-30.
[6] *Biochemical Actions of Hormones* (New York: 1970), ed. Gerald Litwack, Vol. I., p. 466.
[7] R. M. Hare, *The Language of Morals* (Oxford: 1952), pp. 100-101.
[8] *Theories of Ethics* (Oxford: 1967), ed. P. Foot, pp. 78-80. Cf. David Conway, "Description and Evaluation", *Mind* Oct. 1972, pp. 590-94.
[9] Howard T. Milhorn, *The Application of Control Theory to Physiological Systems* (London: 1966), p. 233.
[10] L. E. Bayliss, *Living Control Systems* (San Francisco: 1966), pp. 1-2.
[11] Dettier and Stellar, *Op. cit.* p. 4.
[12] John Searle, *Speech Acts* (Cambridge: 1969), p. 152.
[13] Quoted by Judith P. Swazey in *Reflexes and Motor Integration: Sherrington's Concept of Integrative Action* (Cambridge: 1969), p. 109.
[14] R. Granit, *Charles Scott Sherrington: An Appraisal* (London: 1966), pp. 37-38.
[15] E. C. Tolman, *Purposive Behavior in Animals and Men* (New York: 1932), p. 12.