1980

The Concept of Consciousness

Joseph Margolis
Temple University

Follow this and additional works at: http://digitalcommons.brockport.edu/phil_ex

Part of the Philosophy Commons

Repository Citation
The Concept of Consciousness

by

Joseph Margolis

The key to the problem of consciousness depends on the obvious fact that ascriptions of consciousness are introduced in significantly different, but conceptually related ways, with respect to human persons and to creatures of non-human species or to humans (infants) at phases of development antecedent to their functioning as fully competent persons. The pivotal fact, however, is a naive and simple one: among persons, that is, among ourselves, consciousness is introduced with the admission or recognition of a reporting role shared and understood by the members of a linguistic community. The very difficulty that one has in stating this fact in a non-question-begging way testifies to its fundamental importance. On the other hand, consciousness is introduced in speaking of creatures or systems, or of creatures or systems at certain phases of development below the level of human linguistic ability—or at least recognizably alien to such ability and its manifestations—by way of an observer’s attempt to explain the behavior and existence of such creatures and systems. Consciousness, therefore, is a theoretical posit of some sort with regard to creatures other than human and with regard to the infant phases of human life prior to the development of linguistic ability and the appearance of actual linguistic performance. It serves an explanatory role with regard to human persons as well. But it cannot be introduced in a merely explanatory way, because the very effort to explain and understand the sense in which we are conscious entails our being such. We must, in our own case, introduce consciousness reflexively, that is, as ineliminably entailed by our ability to report and share our thoughts, perceptions, feelings, intentions, and the like. From this point of view, human persons serve as irreplaceable paradigms of what it is to be conscious. Man may not be the center of the universe, but he is the center of every effort to understand it.

To say this much is, of course, not to say what we mean by consciousness. It is only to say that, in a certain plain sense, we cannot deny consciousness to ourselves. This is perhaps a way of recovering the Cartesian cogito without the dubious and puzzling certainty that Descartes assigned to it; the ascription and ascribability of doubt and similar attributes, both reflexively and to others, entails and presupposes, respectively, some form of consciousness. Indeed, the thesis is tautological. It is also vulnerable, in a sense. It is possible, for example, to entertain the elimination of consciousness at both the explanatory and reporting levels—always recognizing that explanation, construed as an activity, has a clear linkage with the reporting use of language. B.F. Skinner’s attempt at a radical behaviorism sought to eliminate, at the level of explaining the molar behavior of humans and other organisms, all use of mental and intentional predicates (except perhaps as a shortcut for the replacing vocabulary). But then, precisely, Skinner may be taken to have intended to provide an analysis of consciousness—in certain non-mental and non-intentional ways—rather than to insure its elimination. In a more fundamental way, W.V. Quine attempted to sketch the general replacement of the intensional features of intensional discourse by an idiom that behaves fully extensionally; but there is no sense in Quine’s program of any need to talk of eliminating mental phenomena—a fortiori, consciousness. Wilfrid Sellars attempted to construct an account of persons solely in terms of roles somehow added to selected aggregates of microtheoretical entities.
adequate for the explanation of all phenomena. Here, though he does not eliminate consciousness altogether, Sellars seems to deny the reality of conscious states—treats them ascriptively (or “forensically”) only. Bolder efforts actually to repudiate consciousness and mental states altogether have never been more than programmatic. Two of the principal lines of exploration have been sketched by Richard Rorty and Paul Feyerabend. But neither has come to grips with the reporting use of language and with the reporting aspect of efforts at scientific explanation itself. In a word, we understand what it would be like to attempt, reductively, to eliminate all reference to consciousness and mental states, but no viable program for their actual elimination has as yet surfaced. There is reason to believe that the difficulties to be overcome are not merely technical but are difficulties both of principle and of “real time” limitations.

The study of consciousness, then, begins reflexively. This in itself has surprisingly powerful consequences, which we may note very briefly. First of all, it can be shown that the analysis of the mental states of animals is inherently anthropomorphized, since the relevant categories formed for the explanation of animal behavior and existence will be conceptually parasitic on just those categories that jointly facilitate the reporting abilities of human beings and the reflexive explanation of their characteristic behavior. Secondly, it can be shown that, since the paradigms of conscious life are supplied by the ability of human persons to report and understand one’s thoughts and feelings, all ascriptions of consciousness and mental states are linguistically modeled even for creatures lacking language. This, of course, is a corollary of the anthropomorphized nature of animal studies. Thirdly, no methodological skepticism (as distinct from error) is possible with regard to the correctness of our description and explanation of animal minds, simply because animals that are not linguistically competent share no reporting function among themselves and pursue no efforts at a theoretical understanding of themselves or of others. Thomas Nagel’s otherwise interesting query, “What is it like to be a bat?” is, therefore, off the mark. Nagel raised the question in order to throw into doubt both a reductive materialism that sought to eliminate interior states of consciousness and an anthropocentric account of consciousness that failed to admit the possibility of alien forms of life. But Nagel’s question is the wrong one, for only human persons (on our present views of the population of the universe) could answer his question. He might have asked, “What is it like to be a dolphin?” if he supposed that we have as yet failed to crack the dolphin language; or, “What is it like to be a Martian?” which only Martians (on the assumption) could at the present time answer.

In virtue of these considerations, one sees that the theory of consciousness is bound to be unified—ranging over humans and animals—precisely because of the methodological differences between how consciousness is introduced in the human context and in the animal; for those differences confirm the priority of reflexive characterizations and the analogical extension of relevant predicates to a range of creatures lacking just those attributes in virtue of which ascriptions of self-consciousness are possible. In fact, we may in general restrict self-consciousness to language-using creatures, that is, to creatures capable of reflexive reference and reflexive predication because of their mastery of language. We may perhaps allow a certain tolerance here, should it be needed, in the direction of incipient linguistic ability among chimpanzees, where self-reference seems quite possible, or even among domesticated dogs, where training seems to invoke responses incipiently like shame and guilt. Self-consciousness, therefore, is conceptually linked to the admission of a reporting role (or, more generously, to behavior taken to be functionally equivalent to actual reporting). Consciousness, on the other hand, is attri-
but able to languageless animals on the strength of some advantage of an explanatory sort, in virtue of which theories that avoid mental states are judged deficient. Hence, the admission of ascriptions of consciousness to animals is tantamount to the rejection of various forms of reductive materialism. Such ascriptions are not so much the adoption of a particular theory regarding the explanation of animal life (though that, doubtless, is their motivation); they entail rather the adoption of a conceptual orientation in which theories of a certain sort (centered on whatever consciousness presupposes or entails) will be favored. Speaking loosely, the admission of consciousness is tantamount to subscribing to a meta-theory or model of alternative theories of a certain range, not to any particular theory. The reason for this distinction will be pursued very shortly.

But here, we must detour briefly to preclude several possible sources of misunderstanding. Human beings are said to speak natural languages. There is in fact no evidence of any human stock that lacks a language (and for every known language, there are bilinguals). By a natural language, we understand a determinate language that is learned in infancy through interaction with the members of a society who actually use that language in their own transactions. However quarrelsome his speculations about an innate grammar may be, even Noam Chomsky does not deny that one cannot speak merely in accord with his linguistic universals; one must actually learn a culturally determinate language—with respect to which, on Chomsky’s theory, our innate linguistic competence manifests itself in terms of linguistic performance. Now, it is quite possible that a machine be programmed to perform linguistically. One can imagine for instance very sophisticated programs being used by the telephone company to “answer” questions about subscribers’ numbers and the like. It is, therefore, quite possible to extend the use of terms like “speak,” “inform,” “give information about,” “answer,” “tell,” and the like, in virtue of which we would be willing to say of a machine that it spoke though it was not capable of consciousness or conscious states. In the human case, the conceptual linkage between speech and consciousness cannot be broken, because the human being learns a natural language and uses it repetitively in just those ways that, paradigmatically, manifest consciousness. It is because machines (on the hypothesis) do not learn language “naturally” from some prelinguistic condition (though they may be said to “learn” to use their language more effectively, with sustained use) that we can entertain the prospect of extending the application of predicates regarding the use of language independently of the application of predicates regarding the manifestation of consciousness. It remains true, nevertheless, that the language imputed to such machines is a language, only on the recognition or interpretation of some human person. So the ascription of linguistic ability to machines is itself conceptually dependent on the admission of a natural language. Furthermore, this conceptual dependence affects every effort we may make to read back, on the strength of machine performance, some interpretation of human performance. The notions of intentionality and information, for instance, are, in the relevant sense, first introduced in the human context and then applied (in different ways) to animals and machines. There is no way known in which it could coherently be claimed that such notions may be defined first for machines and then applied to humans. We see, therefore, that if we treat persons as creatures or systems capable of using language (and of whatever abilities that ability facilitates), artificial persons need not (so far forth) be conscious, though they may be; but natural persons cannot but be conscious. The least reflection, however, shows that related distinctions will be required with respect to the concept of learning itself. “Learning” is an extremely elastic term, often applied to chess-playing machines, for instance, that are
thought not to be conscious; or even to paramecia, whose “behavior” appears to be subject to conditioning—of such an extraordinary sort that the putative “memory” of trained paramecia can be passed on to other untrained specimens merely by ingesting the ground-up remains of their predecessors! But again, since linguistically incompetent animals are ascribed consciousness only at the level of our explanations of their particular form of life, it is entirely possible that the predicate “learning” and similar predicates may be extended to range over phenomena in which no assumptions of consciousness obtain. This seems, for instance, to be intended (with some tendentiousness and conceptual uneasiness, it should be said) by the ethologists.13 We may also notice that, on the argument advanced, there need be no incompatibility between characterizing a system as an automaton and as conscious. Normally, because of the nature of actual machines and actual persons and animals, we are inclined to believe that the one excludes the other. But it is not conceivable (even if it is false) that human life should ultimately be explained on the basis of automaton theory; and it is quite conceivable that the complexity of future machines may justify the ascription of consciousness.14

It must also be emphasized that, since the ascription of consciousness among human persons is reflexively based on the reportorial capacities of humans, it is to be realistically construed. That is, human persons actually possess the attributes of speaking and thinking and the like. No ontology that admits our initial concession regarding the dual nature of ascriptions of consciousness could consistently deny the reality of mental states—unless some ulterior form of eliminative materialism (or surrogate) could be vindicated. Realism with regard to mental states is, then, at least conditionally favored. On the other hand, with respect to animals and machines, it is quite possible to preserve an option between realist and heuristic ascriptions of consciousness and mental states. The issue is rather more complicated than may appear by positing just this contrast, and we shall clarify some of its aspects shortly. By “heuristic,” we here mean that the ascriptions are not realistic; the systems involved do not really or actually or literally have the properties ascribed; only by a façon de parler, a metaphor or analogy, or convenience of reference, do we identify the actual properties involved as if they were genuine forms of consciousness. On this view, by definition, a natural person (that is, a linguistically able system that has learned a natural language as a natural language) must actually be conscious and possess determinate mental states. This is why, for instance, Sellars’ forensic solution of the ontic status of linguistically informed mental states of (human) persons is untenable. To admit the forms of language is to admit the actual psychological capacities of human beings to perform linguistically.

Nevertheless, although animals lacking language can (on the sketch of the theory here advanced) be ascribed consciousness and mental states only in a way that is linguistically modeled, we are not thereby driven to hold that those ascriptions are merely heuristic. On the contrary, it is entirely reasonable to claim that at least the higher mammals actually possess mental states of some sort. The form of the ascription of mental states to animals is heuristic, in the sense, precisely, that the ascriptions are linguistically modeled; but the mental states themselves are realistically attributed to the higher animals. This is simply the consequence of the way in which we have answered Nagel’s question. We alone make ascriptions of consciousness to the bat (if we choose to).15 Doing so, we depend on analogies drawn between non-linguistically informed behavior of bats and the linguistically informed behavior of persons, and we model the ascriptions of the one on what is literally true of the other. Thus, ignoring for the time being the complications that result from making such ascriptions to animals—particularly with respect to puzzles about...
intensionality—the ascriptions may be realistically affirmed though heuristically modeled. There need be no contradiction in this. In fact, it is required if animal psychology be admissible at all. It should perhaps be mentioned, in all fairness, that there have been attempts to model the mental states of animals and humans non-linguistically. But these—notably, for instance, D.M. Armstrong’s effort—have been obliged to introduce notions like “concepts” and “ideas,” which, it may be argued, can be shown not to have eliminated at all an essential appeal to a linguistic model. (We shall return to the issue.) In any case, these considerations are sufficient to dismiss Daniel Dennett’s claim that “a particular thing is an intentional system only in relation to the strategies of someone who is trying to explain and predict its behavior.”

First of all, Dennett construes intentionality in a purely heuristic way. But realism with respect to the mental states of languageless animals is at least eligible, if not well-nigh impossible to deny; also, the mental states of humans cannot all be construed in terms of (linguistically informed) explanatory efforts. Secondly, Dennett confuses intensionality with inten­tionality, holding that the former (in Brentano’s sense) “is primarily a feature of linguistic entities—idioms, contexts;” in fact, he equates the two notions.

Now, it is also conceivable that our very model of linguistic reporting and our explanation of the mental life of languageless animals betray us in some sense—almost “ideologically”—into thinking that the phenomena of either or both animal and human life are correctly analyzed in terms of that model. For example, both ways of speaking emphasize the molar unity of the agent of speech and mental states, and this may be construed as the result of an understandable bias of a certain reflexive habit. Some theorists (notably, structuralists like Claude Lévi-Strauss) actually seem to hold that the underlying structures of human existence are somehow “there” in the external world-independent of the molar bias and psychological capacities of reflexively competent human agents—in spite of the fact that those structures are themselves distinctly intentional in nature. There is reason, therefore, to think that such an approach is either ultimately incoherent or else an extremely attenuated (and somewhat irrelevant) form of caution regarding the finality of any explanatory scheme (including Lévi-Strauss’) or regarding the inescapability of yielding to the tendentious historical currents of one’s own time. If so, it may be safely dismissed. The point of mentioning it, however, is to suggest a certain closure regarding our question, since we should then have considered both infra- and supra-human alternatives to the central role of the mental states of human persons.

It is also possible to attempt to replace the molar agents of speech and mental states with a community of “sub-personal” homunculi, rather in the manner in which Dennett has worked. But, apart from the difficulties of explaining how the apparently molar person, or agent, is completely replaced, without remainder, by such a molecular community, the characterization of such homunculi is quite explicitly and frankly dependent on the idiom suited to the reporting role of human persons. Certain entities, it is true—thoughts and pains and the like—are denied actual existence; but that is hardly equivalent to denying that the having of certain attributes really obtains at the personal level. Only if persons themselves are replaced by molecular homunculi, can the realist ascription to persons (of mental states and consciousness) be denied; but then, the question remains whether, with whatever adjustment may be required, ascriptions to the members of the molecular community will not, in their turn, be tantamount to the admission of consciousness and mental states. The matter will have to be considered more carefully. In any case, the structuralist alternative threatens to remain a mystery, since the very conditions
The Concept of Consciousness

on which its validity depends seem, on the theory, to be inaccessible. And Dennett's alternative is one whose force depends on appraising just how successfully it can replace the intuitively familiar model of persons and of other molar cognitive agents. We should, therefore, turn to consider what it is we claim in claiming that a system is conscious.

What we have, in effect, already conceded is that at least a very large range of attributions normally reserved for conscious persons and conscious animals may, without conceptual difficulty, be ascribed to machines that are not thought to be conscious. We have conceded that this holds, for instance, for predicates regarding speech, providing information, responding to questions; and also, for learning, acquisition of information, and adjustments in behavior dependent on learning and new information. It looks as if a similar extension could be justified for all the verbs of propositional attitude: machines are said to calculate, to solve, to decide, to test, to conjecture, and the like. Possibly the single recalcitrant range of predicates in which, barring consciousness, we should not be inclined to permit the extension covers bodily sensations, mental images, "raw feels," and the like. These are notoriously slighted in Gilbert Ryle's attempt to segregate the "categories" of mental and the physical, so as to preclude interaction.23 They are also stressed in Herbert Feigl's notable effort to vindicate some form of reductive materialism, as providing the most difficult conceptual barrier to any reductive program.24 They are also said to be able to be simlated in a functional or informational sense by machine systems that actually lack such phenomenal episodes.25 And, most recently, they have been reaffirmed (without argument, it must be said), by Karl Popper and John Eccles, as obliging us to return to a Cartesian dualism or a dualism of some as yet unspecified form more or less hospitable to Cartesianism.26

There is a clear sense in which, for instance, the having of pain is inadmissible except on the assumption of consciousness. Much the same seems to be true of ascriptions of images, raw feels, and the like. If the argument rested with such ascriptions only, the matter would be rather quickly resolved, though what the proper analysis of such phenomena might be would still remain an issue of some contention. The trouble is that (i) the ascription of the relevant phenomenal states is paradigmatically managed by way of the reflexive reporting ability of human persons; and (ii) it may well be conceptually possible (and even empirically viable) that organisms or systems be (and be judged) to be conscious, though they lack all such phenomenal states. Hence, bodily sensations and the like are inextricably linked, in a conceptual way, with ascriptions of states of consciousness that need not be characterized by means of any similar phenomenal predicates; and the phenomena themselves are not demonstrably necessary to systems said to exhibit consciousness. Hence, disputes about internal phenomenal states are, largely, rather inaptly linked to the issue of analyzing the nature of consciousness itself.

These considerations confirm that ascriptions of consciousness cannot convincingly be grounded by means of the use of any particular mental predicate: there is, in every individual case, the risk of an extended or attenuated or metaphorical or functionally equivalent use that actually precludes, or ignores, consciousness. The only reasonable supportive strategy, under the circumstances, is to link ascriptions of consciousness to the use of a system of related predicates under at least certain empirically assignable conditions. That system must apply initially to the reportorial performances of human persons and must be entailed in the explanation of languageless creatures said to be conscious. The properties of such a system are so distinctive that invoking it at the explanatory level may be seen at once to challenge the usual forms of reductive materialism. We must, therefore, for the
sake of a fair presentation, consider in turn: (a) the conceptual features of the system; (b) the methodological defense of invoking such a system; and (c) the empirical criteria on which such a system is actually invoked.

We have neutralized the extension of the relevant predicates to nonconscious machines by noticing that their use with respect to machines requires an interpretation on the part of fully conscious persons. This is not simply because all ascriptions are made by conscious persons; on the contrary, we have conceded that they may be made by machines as well. It is rather because the extension of the relevant predicates entails a comparison between applying them to nonconscious systems and to the paradigmatic conscious instances that are ourselves. Ascriptions of mental states, therefore, need not be coextensive with ascriptions of consciousness, but they are conceptually dependent on ascriptions of consciousness nevertheless. For, first of all, mental ascriptions are initially made of human persons capable of reporting and explaining their own mental states, which, paradigmatically, manifest what we mean by consciousness. Secondly, conscious mental states are theoretically ascribed to animals and human infants (at least initially) by human persons, on the grounds of explanatory advantage, in spite of their lack of language. Thirdly, mental states, or analogues of mental states, are ascribed by extension to nonconscious machines—again by human persons gauging functional similarities between man and machines. And finally, unconscious mental states are ascribable only to systems capable in principle of conscious mental states. But, as we have already noted, the system of predicates to be applied to human persons, languageless animals, and prelinguistic infants will be essentially the same—with due provision for the differences noted. Once we grant this much, we may concede as well that the system will, by extended use, be applicable also to certain nonconscious machines. Hence, we must distinguish carefully between features of the system necessary for ascriptions of consciousness (though not sufficient) and those features that serve as criteria for the actual ascription of consciousness itself. There is a characteristic informality about the latter.

The most salient—perhaps the most inclusive—necessary condition for ascriptions of consciousness may be titled rationality. By rationality, we mean a model of the coherence that must hold among the internal mental states of a system and its behavior. Here, a considerable number of distinctions are called for. For instance, we must sort out nested schemata of rationality invoked in making ascriptions to human persons, that range from the most generous species-specific characteristics to more determinate, historically plural, cultural specifications to even more determinate, idiosyncratic, and highly variable personal specifications. The rationality of living creatures varies from species to species; but only in the case of human persons will there appear as well a wide range of alternative schemata of cultural, sub-cultural, and personal models of rationality. The most generic, essentially determinable constraints concern interrelationships among such elements as wants, desires, beliefs, perceptions, intentions, and actions. Thus, ascribing an intention to a human agent normally entails ascribing a belief that what is intended does not yet obtain and that what is intended falls within the agent's power to achieve. Again, ascriptions of desires or wants normally entail ascribing a general congruence between intentions and actions, informed by intervening perceptions and beliefs, and such desires or wants. In short, the ascription of mental states and consciousness conforms with the viability of the species; the pattern of ascription is basically the same for humans and for creatures of other species, except that the mode of life of the different species varies. Thus that the great cats are carnivores must be built into our picture of the species-specific rationality of such creatures. In the human case, our schema normally accommodates some selection of the so-
The Concept of Consciousness

called prudential interests of man, which bridge the animal and personal interests of human beings.27

One important consequence of admitting a species-specific model of rationality is that it becomes logically impossible to ascribe a particular mental state or cognitively informed action without invoking the model in support of other suitably congruent ascriptions of mental states and actions. Random members of a species cannot be said to perceive, for instance, unless they can also be said to want or desire, and to be able to act, or to intend to try to act, in accord with their wants or desires, and their perceptions and beliefs. Variations and idiosyncratic limitations may appear in the lives of particular individuals (for instance, among the insane), but the model is invoked to explain the viability of the species itself. In short, the ascription of particular mental states and particular actions is intelligible only on the assumption that they fit congruously among an entire network of related states and actions. There are no atomic perceptions or beliefs. To determine that a mental state or psychologically informed action obtains is to view the life of an organism under the model of some suitably species-specific rationality. Rationality, therefore, presupposes (and, on an explanatory theory, services) the biological viability of a species. (It may also, of course, be construed analogically with respect to the functioning of a machine.) Here, then, is the methodological motivation for introducing rationality: it is an inference to the best model, or scheme, of explanation for given species, not merely an inference to determinate explanations under such a model.

Another consequence is that the irrationality of individual behavior presupposes, and falls within the range of, the minimal rationality of the species. The model requires only a characteristic congruence among belief, desire, action, and the rest. Particular episodes may well be defective; but to think or behave irrationally is to think or to act—which presupposes the minimal, species-specific competence or rationality of the stock. The problem arises primarily—perhaps nearly exclusively—among human beings. Consequently, there is a strong temptation to construe rationality in the human context as conformity to some determinate ideology or doctrine. A particularly obvious illustration is afforded in a recent account by Philip Pettit of so-called “rational man theory.” Pettit confuses the general rationality of man with adherence to the constraints of decision theory. For instance, he holds that “Every human action springs from a desire or set of desires which, in view of the agent’s beliefs, it promises to satisfy”—where “satisfy” is defined in accord with the thesis that “Beliefs and desires lead to action by familiar rational principles which it is the job of decision theory to spell out.”28 Others may equate rationality with utilitarian calculation.29 A Kantian would insist that rationality entails conformity with the Categorical Imperative. But the exposure of ideology as a sort of pretended objective constraint on rationality30 does not yet bear on the grounds for favoring explanations in terms of the coherence of mental states and actions—over those that do not introduce such states and actions at all. In short, there is no model of mental states and actions that does not provide for the characteristic coherence that holds among relevant ascriptions; but to concede that is to concede a model of species-specific rationality. To perceive, for instance, is to make a discrimination of some kind suitable for guiding action with respect to interest and desire. Putatively isolated perceptions utterly unrelated to desire and action would have no place in promoting the viability of the species—and would be unrecognizable, being linked to no empirically detectable advantage. It is in our effort to understand the survival of the mode of life of given species that we first postulate mental states.

A third consequence of our model is that mental states and actions, and their
congruity, are ascribable at the molar level of organismic life (or at the comparable level for machines), not at any molecular level—that is, not at any level at which a sub-system of a person or sentient animal is designated—unless derivatively from molar ascriptions themselves. Persons think, feel, perceive, believe, and act—not their brains or sequences of neurophysiological processes. Here, the condition of rationality—necessary for consciousness itself—must be viewed in terms of intentionality, or information. Despite the uncertainties of Franz Brentano’s recovery of the concept of intentionality, a number of pertinent generalizations may be offered about its use. Intentionality is a sufficient mark either of the mental or psychological or of what may be done or made by psychologically endowed agents. For instance, machines and artworks exhibit intentionality merely as the products of human work and behavior, in spite of (normally) lacking mental states themselves.

Secondly, wherever psychological states are cognitively construed, their intentional characterization requires propositional objects at least. Thus, to perceive (in the cognitive sense), or to believe, is to perceive, or believe, that \( p \), even if nonpropositional objects may also be correctly ascribed. In accord with the reflexive paradigm of human mental life, one may report that one perceives, or believes, this or that; and our theory of the intelligibility of such reporting requires that what is reportable in first-person contexts is, in a public way, at least in general, ascribable in third-person contexts as well.

The propositional content of human perception and belief is, in being reportable, directly conveyed by suitably selected sentences: \( S \) sees, and reports that \( S \) sees, that there is a horse on the hill. The propositional content of his perception and belief is linguistically modeled; but to perceive and believe is not necessarily to exercise any linguistic ability or to be disposed to. Insofar as any creature or system is ascribed cognitively qualified states, propositional content must be assigned such states. This is, in fact, the only way in which the detailed rationality of cognitively competent systems is articulated. Hence, if the higher animals at least are supposed to perceive or desire, or to act in a way that is informed by perception—in any sense that is cognitively apt—then it is impossible to avoid ascribing propositionally qualified mental states to such creatures, with respect to which a suitable species-specific rationality obtains. Propositions we may treat as heuristic entities (in the sense already provided); but the propositional content of the mental states of animals is only heuristically modeled (though those states are real enough) on the reporting practice of reflexive, self-conscious human persons. This goes entirely contrary to those theorists who wish to concede that animals are aware of pain and perception but who deny that they are capable of belief or of any form of thinking. On the thesis here developed, such a claim is incoherent. The very idea that a lion sees an eland in a cognitively relevant way signifies that (roughly) his seeing that an eland is near is a discrimination pertinently linked to his characteristic appetite and behavior (that are similarly propositionally informed). Such propositional ascriptions are functionally assigned on the assumption that the explanation of the lion’s behavior requires a model of species-specific rationality. The propositional content assigned is an artifact of the explanatory theory. In invoking the theory, however, for the sake of the most effective explanation, we cannot avoid attributing propositional content to the mental states of the creatures or systems involved; also, on a realist theory of scientific explanation, we must suppose that languageless animals do possess mental states—and on our best interpretation, those states must exhibit intentionality or “aboutness.”

It is helpful to note that the puzzles of intensional contexts that affect our analysis of human mental states do not affect in the same way our discourse
The Concept of Consciousness

about the mental states of animals. The reason is elementary: animals lack language, and intensionality is a property of linguistic systems. We must, in working out the species-specific model of rationality appropriate to creatures of different kinds, decide what conceptual limitations and discriminations best characterize one species or another. Mice and pigeons are known to be capable of discriminating certain perceptual forms. The behavior of dogs supports the thesis that, functionally, a dog can discriminate the presence of its master; but, even if a dog's master is the mayor of Philadelphia, there is no reason to think that it can distinguish as such the mayor of Philadelphia. Hence, certain intensional constraints are imposed on the cognitive capacities of animals, by way of specifying the concepts that they can "work" with; but languageless animals are not, on the hypothesis, capable of whatever intensional distinctions presuppose linguistic ability itself. In short, the imposition of a model of rationality entails ascriptions of propositional content to a range of mental states; but such ascriptions, relative to species-specific rationality, are constrained by imputed conceptual limitations appropriate to the species. Hence, (i) contra Armstrong, the very admission of concepts is tantamount to the adoption of a strategy of linguistic modeling: and (ii) like propositions (with which they are inextricably bound), concepts may be construed as heuristic entities, invoked solely for purposes of convenience of explanation.

What we see, then, is that the entire apparatus of rationality and propositionally qualified mental states is invoked to explain the life of organisms and similar molar systems. The congruence that must hold if ascriptions of rationality have any point holds initially among the actions and various states of mind of persons and animals. Paradigmatically, it holds of persons capable of reflexive reporting; on a favorable theory entertained by persons, it may be extended to animals whose functional life appears suitably similar to the life that human beings can report. Information, from this point of view—-not in the sense of communication theory—is intentional in nature. It cannot be ascribed at any molecular level within an organismic or molar system, unless derivatively. Dennett summarizes the issue rather neatly: "the information or content an event [has] within the system [it has] for the system as a (biological) whole." The idea is that it is at least the members of different animal species, or human persons, who exhibit rationally ordered mental states and behavior-intentionality and the possession or processing of information or propositions. Hence, it is to them as well that we first ascribe consciousness.

By analogy with what we have already said, however, it is entirely possible to extend the use of the concept of rationality—or to use it figuratively or heuristically among lower animals (conceivably even among plants, certainly among machines), to which we should deny mental states. What this amounts to is constructing the relatively invariant "responsiveness" of such creatures as, for instance, the tick, as exhibiting, by a façon de parler, a form of rationality comparable with that of creatures capable of cognitively qualified mental states. In short, there is a conceptually ordered declension of ascriptions from the fully cognitive to the functional or teleological that utterly lacks cognitive qualification. At the level of human persons, we concede the paradigm of mental states and actions that are (a) conscious, (b) rationally interrelated, (c) cognitively qualified, and (d) linguistically informed. At the level of the higher animals, we are prepared to admit mental states possessing attributes (a)-(c), where the ascriptions are themselves anthropomorphized in the manner described. Such systems are said (as by Charles Taylor) to be fully purposive—and intentional. By various restrictions on these attributes, we may distinguish alternative functional systems. For example, because machines are actually programmed by human persons, they may, as artifacts, be said to
exhibit a characteristic form of rationality. This is simply to say that, on a human interpretation, machines may be ascribed—with or without consciousness—mental, even cognitive states. Whether or not this must be a mere façon de parler, it involves (at least at the present stage of our technology) a noticeable extension of the relevant attributes well beyond the range of our human paradigms, in the absence of biological analogy. To avoid misunderstanding, such systems may be termed artifactual (if purposive), so as to emphasize the conceptual dependence of relevant ascriptions on the originating purposes of human beings—which, of course, is not to deny that machines may introduce propositionally significant innovations that no human agent could achieve in real time. Alternatively put, they are artifactual though purposive: all ascriptions of rationality and of intentional states are completely borrowed, therefore, from their human makers. In fact, we also speak thus where we mean to deny mental or conscious or cognitive states—for example, in speaking of a thermostat, where the intentional idiom is purely figurative.

This is emphatically not true of the higher animals, regarding which only the form of ascription, not the real states ascribed, is borrowed. In the case of the lower animals, either (a) and (c) are coextensive ranges, not in the sense that every instance of sentience is cognitively informed (which is not even true of humans), but in the sense that every mode of sentience provides occasions of cognition; or, non-cognitive forms of sentience (or at least irritatibility of some sort) may be conceded that functionally resemble the rational organization of conscious mental states. Such systems are said (as by Taylor, to be teleological rather than purposive; the ascription to such systems of propositionally qualified states, or of any form of intentionality, is a mere façon de parler. Teleological ascriptions resemble purposive, or psychologically informed, ascriptions because they are also holistic and because they appear to require peculiar lawlike relations linking processes and some end-state that governs such processes. Whether, in fact, the teleological can be replaced by nonteleological causal regularities we need not here examine. But teleological explanation—for instance, as in homeostatic systems—normally exploits a model of information processing. Either the practice employs a mere façon de parler (where non-teleological explanation is adequate, as in the case of the thermostat; or the model is invoked heuristically, by analogy with machines, because, as perhaps in the case of plants and lower animals, we are unable as yet to explain the relevant phenomena otherwise. We introduce the purposive model heuristically, for the sake, say, of predicting the behavior of a system (say a chess-playing machine, or a plant in difficult terrain), without making any assumptions about the details of its physical organization and without imputing to it mental states realistically. The peculiarity of the ascription of information processing, for instance, at the level of DNA, is simply that information in the relevant sense is propositionally qualified and ascribed to systems incapable of conscious, psychological, or cognitive states. In that sense, it must be entirely an artifact of our explanatory efforts.

We are now in a position to say what we mean by consciousness. Consciousness is (1) the state of any paradigm system capable of using language or of influencing behavior by internal states that that system can report; and (2) the state of any non-paradigm system that is suitably analogous to the relevant states of paradigm systems. This will seem an outrageously simple (and therefore, doubtless, false) account. But it is surprisingly supple and plausible, and it can be shown to avoid certain serious difficulties that other contenders cannot readily manage. Dennett, for example, divides the concept of consciousness thus:

$\forall A \forall t (A \text{ is aware that } \phi \text{ at time } t \iff \phi \text{ is the content of the input state of } A^{\prime} \text{ˈspeech center} \text{ at time } t)$
The Concept of Consciousness

(2) A is aware_2 that \( p \) at time \( t \) if and only if \( p \) is the content of an internal event in \( A \) at time \( t \) that is effective in directing current behavior.\(^{43}\)

His intention is to distinguish between reference to "privileged access" (awareness_1) and reference to "control" (awareness_2). He is persuaded that the two notions "wrongly coalesce" in our intuitive grasp of what it is to be conscious of something.\(^{44}\) But there are certain obvious difficulties with his definitions. First of all, (1) and (2) may be satisfied by systems that are not conscious—if we grant what has already been said about the extension of predicates to machines.\(^{45}\) Secondly, the condition that \( p \) (that is, some proposition) be the content of \( A \)'s speech center, or effective in controlling behavior, is intelligible only on the grounds that \( p \) may (or some \( q \) to which \( p \) is dependently linked) be suitably assigned as the content of some mental state; but, on the argument already presented, this cannot be defended unless \( A \) is a paradigm of a conscious system, or a system ascribed mental states by suitable analogy with such a paradigm. Hence, Dennett's account is either circular or questionbegging. Thirdly, the assignment of \( p \) as the relevant content of some internal state or internal event is intended by Dennett to apply to the parts of \( A \) rather than to \( A \) itself; but, on our argument (and Dennett's own, as cited above), the content or information assigned to an internal event within, or state of an internal part of, a system must "be understood as a function of the function within the whole system of that event [or state]."\(^{46}\) In short, assignments of propositional content to the molecular subsystems of persons or sentient animals is intelligible only if such assignments are conceptually related to assignments made to the relevant molar systems themselves.

Dennett seems to agree with this. But his own reductive program requires that he analyze "a person into an organization of subsystems (organs, routines, nerves, faculties, components—even atoms) and [attempt] to explain the behavior of the whole person as the outcome of the interaction of these subsystems."\(^{47}\) Again, there are difficulties. For one thing, his analytic program violates the principle that (in the relevant sense) the function of a subsystem of a person is assignable only as the subfunction of the characteristic functioning of a molar person. Secondly, the program presupposes, but does not demonstrate, that the functioning of a person is not emergent in the senses attacked, for instance, for alternative reductive reasons, by Herbert Feigl and Mario Bunge:\(^{48}\) for, if it were, then the functionalist, and the cognitivist, analytic program would be doomed.\(^{49}\) The instructive error of Dennett's undertaking is that he confuses the subfunctions of a person's (or animal's) conscious functioning with the functioning of the sub-personal parts of a person or animal—into which that person or animal may be completely analyzed. The sub-functions of person's functioning may be assigned informational or propositional content on the condition that appropriate content be assigned to the mental states of that person. But a person cannot be eliminated in favor of a community of molecular homunculi, or sub-personal parts, for at least two reasons: (a) the ascription of propositional content to the parts has no empirical basis except in terms of analyzing the sub-functioning ingredients within the conscious functioning of the molar system; and (b) the ascription of information or propositional content to the parts is an ascription of content inaccessible to the molar system. Dennett himself says: "We have no direct personal access to the structure of contentful events within us."\(^{50}\) In effect, Dennett's model needs to include (against its own purpose) a relationship between sub-personal homunculi (speaking their own language, so to say) and the molar person, who, one way or another, must be in-
formed in his own idiom of what they say. But this is tantamount to admitting that there is no foreseeable strategy by which to eliminate molar persons and animals that (in a non-reductive sense) are conscious. There can be no reduction of persons to sub-personal parts (of any sort) since sub-personal parts are such only relative to the molar system of which they are the parts: their "information" is nothing but what is parcelled out, on a theory, as ingredient in the information, or intentional content, of what may be independently and antecedently ascribed to the mental states of molar persons. The reduction would be possible only if intentional states could be identified with, or replaced by, non-intentional states; or if "information" adequate to explain the behavior of a molar system could be assigned independently to molecular parts without regard to intentional ascriptions made to the molar system itself. Dennett rejects the first as false to the facts, and rejects the second as incoherent. He has, therefore, undermined the very basis for his own proposal.

These considerations show the resilience of our definition of consciousness. The life of paradigm systems--of human persons--exhibits as such the complexity, the variability, the novelty, the range, the versatility, the creativity of what we mean by consciousness. Application of the notion to languageless animals depends empirically, on a suitably favorable range of behavior on the part of such animals that requires, for its explanation, appeal to a model of rationality similar to that which applies to human beings. Hence, it makes no difference in principle where we draw the cutting line between consciousness and its absence--among languageless creatures. It does not depend, there, on a genuine reporting role; it is invoked solely at the explanatory level (though it may be invoked realistically)--for instance, in explaining the peculiarly inventive swimming play of the cetaceans. Conversely, we dampen, for related reasons, the need to invoke consciousness in explaining the behavior of ticks.

Consciousness, then, is empirically ascribed to persons in terms of their characteristic forms of linguistic exchange, and of the mental states imputed to them that influence their behavior and that they are able to report; by extension, it is empirically ascribed to animals whose mode of communication and behavior is sufficiently complex, versatile, inventive, and the like, that it bears favorably comparison with human paradigms. Such ascriptions to such animals depend essentially on our inability to work in an explanatory way without a full-fledged model of rationality. But such ascriptions are not, for that reason, merely heuristic. On the other hand, in understanding ticks, though we may not need to ascribe consciousness in order to understand and predict the details of their form of life, it is quite possible that a heuristic use of the same model may still remain the most convenient, effective, and accurately predictive instrument to guide our speculations. In the latter case, it remains a façon de parler though it has its (heuristic) predictive use; in the former, it remains realistic, though it is invoked on the strength of explanatory advantage. Explanations of the life of persons, then, require reference to whatever properly affects persons in their cultural milieu--notably, the reasons that they consciously hold that influence them to act as they do. It is true that these accounts must be supplemented by attention to processes that operate at molecular levels within persons--in terms, for instance, of perception, inference, memory storage, retrieval, and the like. But causal factors at such levels, insofar as they are assigned informational or intentional content, are assigned such content only congruently with whatever may be ascribed within the repertory of mental states assignable to molar persons. This at least is what is required if persons and their linguistically informed mental states are genuinely emergent phenomena. For here, two essential constraints must be conceded. In the first place, intentional "content cannot be
The Concept of Consciousness

described, only assigned, on a model of rationality, to appropriate systems. And in the second, there is no known way in principle to replace our intentional models (once introduced) by non-intentional ones. But the paradigm application of intentional models is the conscious life of human persons.

FOOTNOTES

8 Thomas Nagel, "What is it like to be a bat?," Philosophical Review, LXXXIII (1974).
12 Margolis, loc. cit.
16 See Margolis, loc. cit.
19 Ibid., p. 3.


30 Cf. Habermas, loc.cit.


36 See further Margolis, *Persons and Minds*.

37 Cf. Sayre, loc. cit.

38 Dennett, *Content and Consciousness*, Ch. 4.
The Concept of Consciousness


44 *Brainstorms*, p. 31.


46 Ibid., p. 163.


49 Dennett actually says: "So far as I can see, however, every cognitivist theory currently defended or envisaged, functionalist or not, is a theory of the sub-personal level. It is not at all clear to me, indeed, how a psychological theory-as distinct from a philosophical theory-could fail to be a sub-personal theory. So the functionalists' problem of capturing the person as subject of experience must arise as well for these cognitivist theories," *Brainstorms*, pp. 153-154.

50 Ibid., p. 169.


53 Dennett, *Content and Consciousness*, p. 83.

54 Dennett declares much too sanguinely (but pertinently for his own program) that "... although no neat synonymy or correlation between intentional and non-intentional sentences has been discovered or proposed, sense has been made of the lesser claim that certain types of physical entities are systems such that their operations are naturally to be described in the Intentional mode-and this, only in virtue ultimately of their physical organization. The force of 'naturally' here is this: although such systems are ultimately amenable to an extensional theory of their operations, their outward manifestations are such that they can be intelligibly described at this time, within our present conceptual scheme, only in the Intentional mode," ibid., p. 89. A full account of the prospects of the intended replacement is provided in Margolis, *Persons and Minds*. 

Published by Digital Commons @Brockport, 1980