

INTENTIONALITY AND ITS PLACE IN NATURE

1.

Intentionality is that feature of certain mental states and events that consists in their (in a special sense of these words) being *directed at*, being *about*, being *of*, or *representing* certain other entities and states of affairs. If, for example, Robert has the belief that Ronald Reagan is President, then his belief is an intentional state because in the appropriate sense his belief is directed at, or about, or of, or represents Ronald Reagan and the state of affairs that Ronald Reagan is President. In such a case Ronald Reagan is the *intentional object* of Robert's belief, and the existence of the state of affairs that Ronald Reagan is President is the *condition of satisfaction* of his belief. If there is not anything that a belief is about, then it does not have an intentional object; and if the state of affairs it represents does not obtain, it is not satisfied.

Ascriptions of intentionality are of differing kinds, and as these differences have been a source of confusion, I will begin by sorting out some of them. Consider the statements made in utterances of the following sentences:

- A. Robert believes that Ronald Reagan is President.
- B. Bill sees that it is snowing.
- C. "Es regnet" means it's raining.
- D. My car thermostat perceives changes in the engine temperature.

Each of these statements ascribes intentionality, but the status of the ascriptions is different. A simply ascribes an intentional mental state, a belief, to a person; B does more than that, since to say of someone that he sees that something is the case implies not only that he has a certain form of intentionality but also that the intentionality is satisfied, i.e., that the conditions of satisfaction actually obtain. "See", like "know" but unlike "believe", is a success verb: *x sees that p* entails *p*. There is an intentional phenomenon reported by B, since in order to see something

one has to have a visual experience, and the visual experience is the vehicle of intentionality. But B does more than report a visual experience; it also reports that it is satisfied. Furthermore, visual experiences differ from beliefs in being conscious mental events rather than states. A man asleep can be literally said to believe that such and such, but not to see that such and such. Both beliefs and visual experiences are *intrinsic* intentional phenomena in the minds/brains of agents. To say that they are intrinsic is just to say that the states and events really exist in the minds/brains of agents; the ascription of these states and events is to be taken literally, not just as a manner of speaking, nor as shorthand for a statement describing some more complex set of events and relations going on outside the agents.

In this last respect the ascription of intentionality in A and B differs from C and D. C literally ascribes intentionality, though the intentionality is *not intrinsic*, but *derived*. It is literally true that the sentence "Es regnet" means it's raining, but the intentionality in question is not intrinsic to the sentence. That very sentence might have meant something else or nothing at all. To ascribe this form of intentionality to it is shorthand for some statement or statements to the effect that speakers of German use the sentence literally to mean one thing rather than another, and the intentionality of the sentence is derived from this more basic form of intrinsic intentionality of speakers of German. In D, on the other hand, there is no literal ascription of intentionality at all because my car thermostat does not literally have any perceptions. D, unlike C, is a metaphorical ascription of intentionality; but, like C, its point depends on some intrinsic intentionality of agents. We use car thermostats to regulate engine temperatures and therefore they must be able to respond to changes in temperature. Hence the metaphor; and hence its harmlessness, provided we don't confuse the analysis of A, B, and C, with that of D.

To summarize: even from this short list of statements there emerge several distinctions, which – in addition to the usual distinction between conscious and unconscious forms of intentionality – we will need to keep in mind.

1. The distinction between ascriptions of intentionality that imply that the intentional phenomenon is satisfied and those that do not, as illustrated by A and B.
2. The distinction between intentional states and intentional

events, as also illustrated by A and B. (For brevity, I will sometimes call both “intentional states”.)

3. The distinction between intrinsic intentionality and derived intentionality, as illustrated by the distinction between the intentionality ascribed in A and B, on the one hand, and C on the other.
4. The distinction between literal ascriptions of intentionality such as A, B, and C, whose truth depends on the existence of some intentional phenomenon, whether intrinsic or derived; and metaphorical ascriptions, such as D, which do not literally ascribe any intentionality at all, even though the point of the metaphorical ascription may depend on some intrinsic intentionality of human agents.

In the rest of this paper I will deal only with intrinsic intentionality and the question I will discuss, to put it broadly is as follows: What is the place of intrinsic intentionality in nature?

2.

Intentional mental phenomena are part of our natural biological life history. Feeling thirsty, having visual experiences, having desires, fears, and expectations, are all as much a part of a person’s biological life history as breathing, digesting, and sleeping. Intentional phenomena, like other biological phenomena, are real intrinsic features of certain biological organisms, in the same way that mytosis, meiosis, and the secretion of bile are real intrinsic features of certain biological organisms.

Intrinsic intentional phenomena are caused by neurophysiological processes going on in the brain, and they occur in and are realized in the structure of the brain. We do not know much about the details of how such things as neuron firings at synapses cause visual experiences and sensations of thirst; but we are not totally ignorant, and in the cases of these two intentional phenomena we even have pretty good evidence about their locations in the brain. That is, for at least some intentional phenomena, we have some idea of the special role of certain brain organs, such as the visual cortex or the hypothalamus, in producing the intentional phenomena. More important for our present discussion, our ignorance of how it all works in the brain is an empirical ignorance of

details and not the result of a metaphysical gulf between two incommensurable categories, the "Mind" and the "Body", which would prevent us from ever overcoming our ignorance. Indeed, the general sorts of relations involved between mental phenomena and the brain are quite familiar to us from other parts of nature. It is common in nature to discover that higher level features of a system are caused by the behavior of lower level microentities and realized in the structure of the system of microentities. For example, the solidity of the metal in the typewriter I am currently hammering on is caused by the behavior of the microparticles that compose the metal, and the solidity is realized in the structure of the system of microparticles, the atoms and molecules. The solidity is a feature of the system but not of any individual particle. Analogously, from what we know about the brain, mental states are features of the brain that are caused by the behavior of the elements at the microlevel and realized in the structure of the system of microelements, the neurons. A mental state is a feature of the system of neurons but not of any particular neuron. Furthermore, on this account there is no more reason for counting mental states as epiphenomenal than there would be for counting any other intrinsic, higher level features of the world, such as the solidity of this typewriter, as epiphenomenal.

In sum, certain organisms have intrinsic intentional states, these are caused by processes in the nervous systems of these organisms, and they are realized in the structure of these nervous systems. These claims should be understood in as naturalistic a sense as the claims that certain biological organisms digest food, that digestion is caused by processes in the digestive tract, and that it all goes on in the stomach and the rest of the digestive tract. Part of our difficulty in hearing the former claims naturalistically derives from the fact that the traditional vocabulary for discussing these problems is designed around a seventeenth century conception of the "mind/body problem". If we insisted on using the traditional jargon we might say: monism is perfectly compatible with dualism, provided it is a property dualism; and property dualism is compatible with complete physicalism, provided that we recognize that mental properties are just one kind of higher level property along with any number of other kinds. The view is not so much dualism as polyism, and it has the consequence that intrinsic mental properties are just one kind of higher level physical property among many others (which is perhaps a good reason for not using the traditional jargon at all).

It is a remarkable fact about contemporary intellectual life that the

existence of intrinsic intentional phenomena is frequently denied. It is sometimes said that the mind with its intentional states is something abstract, such as a computer program or a flow chart; or that mental states have no intrinsic *mental* status because they can be entirely defined in terms of their causes and effects; or that there aren't any *intrinsic* mental states, but rather that talk about mental states is just a manner of speaking that enables us to cope with our environment; and it is even sometimes said that mental terms should not be thought of as standing for actual things in the world at all. To catalogue the reasons that people have had for holding these views and denying everything that biology has to tell us about the brain would be to catalogue some of the leading intellectual confusions of the epoch. One, though only one, of the sources of confusion is the deeply held belief that if we grant the existence of intrinsic intentional states we will be confronted with an insoluble "mind/body" or "mind/brain" problem. But, to paraphrase Darwin, it is no more mysterious that the brain should cause mental phenomena than that bodies should have gravity. One might, and indeed one should, have a sense of awe and mystery in the face of both facts, but that sense would no more justify us in the denial of the existence of mental states than it would justify us in the denial of the existence of gravity.

Some philosophers feel that I am unjustified in simply asserting the existence of intrinsic intentional mental states and events in the world. For, they argue, might not the progress of science show them to be an illusion in the way that the appearance of the sun rising and setting over a stationary earth was shown to be an illusion? Isn't it just as prescientific to believe in intrinsic mental states as it is to believe that the earth is flat and in a fixed position in the universe?¹

But if we confine our attention for the moment to conscious intentional mental events and states – and they are, after all, the primary forms of intentionality – we can see that the analogy between the belief in a flat and fixed earth and the belief in the existence of mental phenomena breaks down. In the case of the earth there is a clear distinction between how things are and how they seem to be, but in the case of the very existence of conscious mental phenomena it is hard to know what a parallel distinction would look like. I know more or less exactly what it means to say that, though the earth seems flat and fixed, it is in fact not flat and fixed but rather is round and mobile; but I haven't the faintest idea what it would mean to say that, though it seems

to me that I am now conscious, in fact I am not really conscious but rather I am . . . What?

The reason we are unable to fill in the gap with anything that does not seem preposterous has been familiar since Descartes: If it consciously seems to me that I am conscious then that is enough for me to be conscious. And that is why there cannot be a general “how things seem”/“how they really are” distinction for the very existence of conscious mental states.

This is not, of course, to say that we cannot discover all sorts of surprising and counterintuitive things about our mental life, about the nature and mechanisms of both conscious and unconscious mental states. But it is to say that the distinction between how things seem and how they really are cannot apply to the *existence* of our own conscious mental phenomena.

3.

Since the resistance to treating consciousness and intentionality naturalistically, as just higher level properties among others, is so pervasive, and since the view of the place of intentionality in nature advanced in this article is so much out of step with what is currently accepted, I want to probe these issues a little more deeply. If one reads the standard literature on the “mind/body problem” over the past thirty years,² since the publication of Ryle’s *The Concept of Mind* (1949), one discovers a curious feature of this continuing dispute. Almost all the participants on both sides tacitly assume that the specifically mental features of conscious mental events cannot be ordinary physical features of the world like any other higher level features of the world. This assumption is often masked from us by the way theses are stated. Thus, when the identity theorist tells us that mental states just *are* brain states, there is a way of hearing that thesis which is perfectly consistent with our common-sense assumption of the intrinsic and irreducible character of consciousness and other forms of intentionality. We can hear the thesis as saying that mental processes are just processes going on in the brain in the way that digestive processes are processes going on in the digestive tract. But in general that is not what identity theorists are claiming. Under close scrutiny of the texts, particularly those parts of the texts where they are replying to dualist adversaries, it turns out that in general identity theorists (materialists, physicalists, functionalists,

etc.) end up by denying the existence of intrinsically mental features of the world. J. J. C. Smart, with typical candor, states the position clearly in responding to J. T. Stevenson:

My reply is that I do not admit that there are any such *P*-properties [i.e. properties of sensations that would prevent us from defining 'sensation' in terms of properties in a physicalist scheme]. (1970, p. 93)

Now why does Smart feel it necessary to deny what seems to Stevenson (and to me) an obvious truth? I believe it can only be because, in common with the tradition since Descartes, he thinks that to grant the reality of conscious mental phenomena is to grant the existence of some mysterious phenomena, some sort of "nomological danglers" beyond the reach of the physical sciences. On the other hand, consider those who challenge the tradition of materialism by asserting such obvious facts as that they are currently having a series of conscious states. They seem to think their claim commits them to some form of dualism, as if in asserting obvious facts about our waking life they are committed to the existence of some ontological category different from that of the ordinary physical world we all live in. One group of philosophers sees itself as defending the progress of science against residual superstitions. The other group sees itself as asserting obvious facts that any moment of introspection will reveal. But both accept the assumption that naive mentalism and naive physicalism must be inconsistent. Both accept the assumption that a purely physical description of the world could not mention any mental entities.

These are false assumptions. Unless one defines "physical" and "mental" in such a way that one is the negation of the other, there is nothing in our ordinary notions of mental phenomena and physical reality that excludes cases of the former from being instances of the latter.

Why then, to continue the investigation a step further, do both sides make this apparently obvious mistake? I think the answer must be that they take very seriously a whole tradition, going back at least to Descartes, with its endless disputes about substance, dualism, interaction, emergence, ontological categories, the freedom of the will, the immortality of the soul, the presuppositions of morality, and the rest of it. And in large measure this tradition revolves around the assumption that "mental" and "physical" name mutually exclusive categories. But suppose for a moment that we could forget all about this entire

tradition. Try to imagine that we are simply investigating the place in nature of our own human and animal mental states, intentional and otherwise, given what we know about biology, chemistry, and physics and what we know from our own experiences about our own mental states. I believe if we could forget the tradition, then the question as to the place of such states in nature would have an obvious answer. They are physical states of certain biochemical systems, viz., brains. But there is nothing reductive or eliminative about this view. Mental states with all their glorious or tiresome features – consciousness, intentionality, subjectivity, joy, anguish and the rest – are exactly as we knew they were all along.

Lest my view be misunderstood, I should like to state it with maximum simplicity. Take the most naive form of mentalism: There really are intrinsic mental states, some conscious, some unconscious; some intentional, some nonintentional. As far as the conscious ones are concerned they pretty much have the mental properties they seem to have, because in general for such properties there is no distinction between how things are and how they seem. Now take the most naive version of physicalism: The world consists entirely of physical particles, including the various sorts of relations between them. As far as real things in the world are concerned there are only physical particles and various arrangements of physical particles. Now, my point is that it is possible to accept both of these views exactly as they stand, without any modification whatever. Indeed the first is simply a special case of the second.

4.

Granted that intentional mental states really do exist and are not to be explained away as some kind of illusion or eliminated by some sort of redefinition, what role do they play in a naturalistic or scientific description of nature?

Just as it is a biological fact that certain sorts of organisms have certain sorts of mental states, so it is equally a biological fact that certain mental states function causally in the interactions between the organism and the rest of nature and in the production of the behavior of the organism. It is just a fact of biology that sometimes thirst will cause an organism to drink water, that hunger will cause it to seek and consume food, and that sexual desire will cause it to copulate. In the

case of human beings, at a much more sophisticated though equally biological level, the beliefs a person has about what is in his or her economic interest may play a causal role in how he or she votes in political elections, literary preferences may play a causal role in the purchase and reading of books, and the desire to be someplace other than where one is may play a causal role in a person's buying a plane ticket, driving a car, or getting on a bus. Though the fact of causal relations involving intentional mental states is pretty obvious, what is a great deal less obvious is the logical structure of the causal relations involved and the consequent implications that those causal relations have for the logical structure of the explanation of human behavior.

Explanations involving intentionality have certain logical features not common to explanations in the other physical sciences. The first of these is the specific role of intentional causation in the production of certain sorts of animal and human behavior. The essential feature of intentional causation is that the intentional state itself functions causally in the production of its own conditions of satisfaction or its conditions of satisfaction function causally in its production. In the one case the representation, as a representation, produces what it represents; in the other case the object or state of affairs represented functions causally in the production of its representation. This point can be made clear by considering some examples. If I now have a strong desire to drink a cup of coffee and I act on that desire so as to satisfy it, then the desire whose content is

(that I drink a cup of coffee)

causes the very state of affairs, that I drink a cup of coffee. Now, in this simple and paradigmatic case of intentional causation, the desire represents the very state of affairs that it causes. The much discussed "internal connection" between "reasons for action" and the actions that they cause is just a reflection of this underlying feature of intentional causation. Since the cause is a representation of that which it causes, the specification of the cause, as cause, is indirectly already a specification of the effect.

Sometimes, indeed, the intentional state has as part of its conditions of satisfaction, as part of its intentional content, that it must function causally if it is to be satisfied. Thus, for example, intentions can only be satisfied if the actions that they represent are caused by the intentions that represent them. In this respect intentions differ from desires: a

desire can be satisfied even if it does not cause the conditions of its satisfaction; whereas an intention can be satisfied only if it causes the rest of its own conditions of satisfaction. For example, if I want to be rich and I become rich, my desire will be satisfied even if that desire played no causal role in my becoming rich; but if I intend to earn a million dollars and I wind up with a million dollars quite by accident, in such a way that my intention to earn the money played no causal role consciously or unconsciously in my getting it, then although the state of affairs represented by my intention came about, the intention itself was not satisfied, i.e., the intention was never carried out. Intentions, unlike desires, have intentional causation built into their intentional structure; they are causally self-referential in the sense that they can only be satisfied *if they cause* the very action they represent. Thus, the prior intention to drink a cup of coffee differs in its content from the desire to drink a cup of coffee, as we can see by contrasting the following representation of the conditions of satisfaction of a prior intention with our representation above of the conditions of satisfaction of the corresponding desire:

(that I drink a cup of coffee and that this prior intention cause that I drink a cup of coffee).

Cases of “volition”, such as desires and intentions, have what I call the “world-to-mind direction of fit” (the aim of the state is to get the world to change to match the content of the desire or intention) and the “mind-to-world direction of causation” (the mental state causes the state of affairs in the world that it represents). Cases of “cognition”, such as perception, memory, and belief, function conversely as far as direction of fit and intentional causation are concerned. Thus, they have the mind-to-world direction of fit (the aim of the mental state is not to create a change in the world, but to match some independently existing reality); and, where intentional causation functions in the production of the intentional state, they have the world-to-mind direction of causation (if I correctly perceive or remember how things are in the world, then their being that way must cause my perceiving them or remembering them as being that way).

5.

I believe that a full account of the role of intentionality and its place in

nature requires much more study of intentional causation than has yet been done or than I can undertake in this essay. But by way of giving the reader some idea of the importance of intentional causation, I want to mention just three of the implications of this brief sketch of intentional causation for a causal account of human and animal behavior and for ways in which such a causal account differs from certain standard models of what we have as canonical explanations in the usual natural sciences.

1. In any causal explanation, the propositional content of the explanation specifies a cause. But in intentional explanations the cause specified is itself an intentional state with its own propositional content. The canonical specification, therefore, of the cause in an intentional explanation doesn't just *specify* the propositional content of the cause, but it must actually *repeat* in the explanation (at least some of) the propositional content that is functioning causally in the operation of the cause. So, for example, if I buy a plane ticket because I want to go to Rome, then in the explanation:

I did it because I want to go to Rome.

I repeat the very propositional content functioning in the operation of the desire:

I want to go to Rome.

Intentional explanations are more or less adequate as they accurately repeat in the explanation the propositional content functioning in the cause itself. It is a further consequence of this feature that the concepts used in the canonical form of the explanation don't just describe a cause; rather, the very concepts themselves must function in the operation of the cause. So, if I say that a man voted for Reagan because he thought it would increase the probability that he would be rich and happy, such concepts as being rich and being happy can be used in the explanation to specify a cause only if they also function as part of the cause.

These features have no analogue in the standard physical sciences. If I explain the rate of acceleration of a falling body in terms of gravitational attraction together with such other forces as friction operating on the body, the propositional content of my explanation makes reference to features of the event such as gravity and friction, but the features themselves are not propositional contents or parts of

propositional contents.

This is a familiar point in the history of discussions of the nature of psychological explanation, but it seems to me that it has not been properly stated or appreciated. I believe it is part of what Dilthey (1962) was driving at when he said that the method of *Verstehen* was essential to the social sciences, and it was part of what Winch (1958) was driving at when he said that concepts used in the explanation of human behavior must also be concepts that are available to the agent whose behavior is being explained. I think an analysis of intentional causation would provide us with a deeper theoretical understanding of the points that Dilthey and Winch were after.

2. Statements of intentional causation do not require the statement of a covering law in order to be validated or in order to be causally explanatory. In a subject like physics we assume that no causal explanation of a phenomenon is fully explanatory unless it can be shown to instantiate some general law or laws. But in the case of intentional causation this is not generally the case. Even if we believe that there are laws, stateable in some terms or other, which any given instance of behavior instantiates, it is not essential to giving a causal explanation of human behavior in terms of intentional causation that we be able to state any such laws or even believe that there are such laws.

3. Teleological forms of explanation are those in which a phenomenon is explained in terms of goals, aims, purposes, intentions, and similar phenomena. If teleological explanation is really a subclass of scientific explanation, it would appear that nature must actually contain teleological phenomena. The account of intentionality and its place in nature that I have been urging has the consequences both that nature contains teleological phenomena and that teleological explanations are the appropriate forms of explanation for certain sorts of events. Indeed, it is an immediate logical consequence of the claim that goals, aims, purposes, and intentions are intrinsic features of certain biological organisms that teleology is an intrinsic part of nature, for by definition such phenomena are teleological. And it is an immediate consequence of the characterization I have given of these phenomena that teleological explanations are the appropriate forms for explaining certain sorts of events, since these phenomena cause events by way of the form of intentional causation that is peculiar to teleology.

All the states I have called "teleological" have the world-to-mind direction of fit and the mind-to-world direction of causation. The

explanatory role of citing such states in teleological explanations can best be illustrated by example. Consider the case of an animal, say a lion, moving in an erratic path through tall grass. The behavior of the lion is explicable by saying that it is stalking a wildebeest, its prey. The stalking behavior is caused by a set of intentional states: it is *hungry*, it *wants* to eat the wildebeest, it *intends* to follow the wildebeest with the *aim* of catching, killing, and eating it. Its intentional states represent possible future states of affairs; they are satisfied only if those states of affairs come to pass (world-to-mind direction of fit); and its behavior is an attempt to bring about those states of affairs (mind-to-world direction of causation). The claim that teleology is part of nature amounts to the claim that certain organisms contain future-directed intentional states with the world-to-mind direction of fit, and that these states are capable of functioning causally to bring about their conditions of satisfaction.

It is worth emphasizing the logical features of teleological explanation because on some accounts a teleological explanation explains an event by the occurrence of a future event, as if, for example, the eating of the prey explained the stalking behavior.³ But on my account this conception has things back to front. All valid teleological explanations are species of explanation in terms of intentional causation, and there is no mysterious backwards operation of intentional causation. The stalking behavior at time t_1 is explained by present and prior intentional states at t_1 and t_0 , and these aim at the eating behavior of t_2 .

In the great scientific revolution of the seventeenth century the rejection of teleology in physics was a liberating step. Again in the great Darwinian revolution of the nineteenth century the rejection of a teleological account of the origins of the species was a liberating step. In the twentieth century there has been an overwhelming temptation to complete the picture by rejecting teleology in the sciences of man. But ironically the liberating move of the past has become constraining and counterproductive in the present. Why? Because it is just a plain fact about human beings that they do have desires, goals, intentions, purposes, aims, and plans, and these play a causal role in the production of their behavior. Those human sciences in which these facts are simply taken for granted, such as economics, have made much greater progress than those branches, such as behavioristic psychology, which have been based on an attempted denial of these facts. Just as it was bad science to treat systems that lack intentionality as if they had it, so it is

equally bad science to treat systems that have intrinsic intentionality as if they lacked it.

NOTES

¹ Rorty (forthcoming), p. 84.

² I am thinking of the sort of articles to be found in Borst (1970), Rosenthal (1971), and Block (1980).

³ For a discussion of this conception see Braithwaite (1953), Chapter X.

REFERENCES

- Block, N. (ed.): 1980, *Readings in Philosophical Psychology*, Vol. 1, Harvard University Press, Cambridge, Mass.
- Borst, C. V. (ed.): 1970, *The Mind-Brain Identity Theory*, Macmillan, London.
- Braithwaite, R. B.: 1953, *Scientific Explanation*, Cambridge University Press, Cambridge, England.
- Dilthey, W.: 1962, *Meaning in History*, H. P. Rickman, (trans. and ed.), New York.
- Rorty, R.: forthcoming, 'Mind as Ineffable', in R. Q. Elvee (ed.), *Mind and Nature*, Harper & Row, San Francisco.
- Rosenthal, D. M. (ed.): 1971, *Materialism and the Mind-Body Problem*, Prentice-Hall, Englewood Cliffs, N.J.
- Ryle, G.: 1949, *The Concept of Mind*, Hutchinson's University Library, London.
- Smart, J. J. C.: 1970, 'Further Remarks on Sensations and Brain Processes', in Borst (1970), pp. 93-94.
- Winch, P.: 1958, *The Idea of a Social Science*, Routledge & Kegan Paul, London.

Department of Philosophy
 University of California, Berkeley
 Berkeley, California 94720
 U.S.A.