Understanding, truth, and explanation

Josep Corbí

In this paper I intend to discuss the relationship between understanding, truth, and explanation, starting from the problem of the interpretation of alien belief systems. The examination of this will lead us to present an analysis of the structure of understanding, and to defend its cognitive universality in so far as any identification of reality reproduces the structure of the process of understanding. This cognitive universality does not necessarily exclude, however, the possibility of reducing understanding to causal relations. With respect to this, I shall discuss the plausibility of the views of Quine, Searle and Davidson on the kind of relation that holds between cognitive processes and the causal relations in nature. In my opinion, Quine's position is untenable; Searle's view, if non-trivially interpreted, is at least implausible; and, finally, Davidson's anomalous monism can be construed in such a way that it is compatible with our analysis of the structure of understanding and with the thesis of its cognitive universality, in spite of Davidson's commitment to the ontological universality of causal relations.

I

In 1935, Evans-Pritchard wrote:

> Scientific notions are those which accord with reality both with regard to the validity of their premises and to the inferences drawn from their propositions. . . . Logical notions are those in which according to the rules of thought inferences would be true were the premises true, the truth of the premises being irrelevant.
>
> A pot has broken during firing. This is probably due to grit. Let us examine the pot and see if this is the cause. That is logical and scientific thought. Sickness is due to witchcraft. A man is sick. Let us consult the oracles to discover who is the witch responsible. That is logical and unscientific thought.[1]

By introducing this double criterion of rationality to evaluate any system of belief, Evans-Pritchard sought to mark himself off from certain

thinkers who interpreted magical beliefs as basically contradictory[2] and who, thereby, attributed an illogical system of thought to certain cultures. Against such a position Evans-Pritchard held that magical beliefs are internally coherent. However he was less generous when it came to evaluating the truth of such beliefs, since he considered the premises upon which magical beliefs are based to be strictly false in so far as they do not match reality. In this way Evans-Pritchard seems to be following the positivist model of rationality, according to which any system of belief is to be evaluated by just two criteria, namely, its internal coherence and its truth, the latter understood as correspondence with reality.

However, just as we are right to feel suspicious about any interpretation which considers a system of thought as fundamentally incoherent, so too we should be loath to accept a standpoint which views magical thought (which has, after all, been adopted by a very considerable proportion of humanity throughout history) as being based on premises which are strictly false. Doubts such as these have led certain philosophers to propose a different strategy when faced with the problem of interpreting magical beliefs. Broadly speaking, we can say that whereas in the past thinkers started from the assumption that the positivist criteria of rationality were universally valid, and went on to disqualify any model of thought which did not fit in with them, today one tends to begin with the assumption that all models of thought are basically rational, and go on from there to determine the specific criteria of rationality behind each model.

This new strategy has emerged from a revision of the positivist analysis of the notion of reality. Thus, in the text of Evans-Pritchard, the word 'reality' seems to be used to refer to some element external to thought and to language with which different beliefs are to be compared in order to determine their truth. Now, if such a comparison is to be possible, we must be able to identify both of its terms independently. So Evans-Pritchard seems to assume that it is possible to identify reality without our identification being biased through its commitment to a particular language or system of thought. Let us examine to what extent such an assumption is valid.

Establishing the identity of objects entails the notion of a rule, since, given the infinite aspects under which we may consider an object, we need a rule to determine precisely those aspects which are to be considered relevant for maintaining the numerical and qualitative identity of a given object. Furthermore, as different communities can follow different rules to establish what is real, we seem to be forced to conclude that in spite of the positivists' claim any identification of reality is conditioned by the system of rules in accordance with which it has been carried out. Consequently one cannot naïvely maintain that the way science characterizes reality is linguistically unconditioned, and so it is not possible to claim that it can be used as the universal criterion of validity for all systems of belief.

This approach seems to lead to conclusions which are intuitively as paradoxical as those which followed from the universalisation of the positivist criteria of rationality, since now we find ourselves with no criteria whatsoever with which to compare the rationality of different systems of belief. Full-blooded relativism is just as suspect as is the assumption that until recently all human thought had basically been a mistake. I shall endeavour to offer an analysis of the interpretation of alien systems of belief, which will enable us to elude both full-blooded relativism and naïve universalism.[3]

Let us first consider what is involved in the process of identifying a system of belief. The idea of a system of belief presupposes the idea of a subject to whom this belief is attributed. Now I think that it can be shown that when attributing beliefs to a subject, one must start from two basic assumptions: a) that the subject's beliefs are basically coherent, and b) that when the subject holds a set of beliefs he is also committed to the belief that it is, in general, rational to hold this set of beliefs.

With respect to (a), it seems clear that it makes no sense to claim that a subject maintains a basically contradictory system of beliefs, since to maintain contradictory beliefs amounts to not maintaining any belief at all or, if one prefers, to maintaining just any belief. Thus, if the beliefs identified as contradictory belonged to the core of a belief system, then the rest of the beliefs would lose their content, and we would no longer be entitled to say that they are beliefs at all. As a result, one *a priori* assumption underlying every process of identification of beliefs is that the subject's beliefs are basically coherent. Whence it follows that any interpretation which presents a system of belief as basically self-contradictory is, in fact, an interpretation that has identified the contents of the beliefs incorrectly.[4]

As regards (b), we must bear in mind that in saying that a subject believes something, we are saying that the subject believes something about reality. Now, believing something about reality implies believing that a statement is true; but holding a statement to be true implies the ability, on the part of the subject, to distinguish between evidence that will back up, and evidence that will question, his commitment to his belief. It follows from this that a subject's commitment to the truth rather than to the falsehood of a statement cannot be, in the majority of cases, arbitrary. The subject must have some reason however vague to adopt a belief and, moreover, the subject must be able to consider new reasons which could lead him to modify his commitment to the truth of the statement in question. Therefore one cannot determine the reasons for holding a belief without discriminating between what constitutes and what does not constitute a reason for believing something. But this is the same as distinguishing between good reasons and bad reasons for believing something. The idea of a reason for believing is intrinsically normative.

Nevertheless, this normativity commits not only the subject who is being interpreted, but the interpreter himself, since in the process of identifying beliefs he must be committed to the statement that certain reasons do justify certain beliefs.

Against this line of argument one could object that it relies on an excessively rationalist conception of the adoption of beliefs and that, in fact, most of the beliefs held by a subject are merely the result of a process of socialization, and so do not obey any type of reason but are due to social conditioning. There is a sense in which I believe that this objection is correct, since I do not wish to deny the influence of the social or cultural context in which an individual develops on the beliefs he holds to be true. However, this conditioning does not necessarily alter the fact that, in general, the subject must consider it rational to adopt the beliefs that, conditioned by his cultural context, he does in fact adopt. It could be argued that in some cases we can meaningfully say that a given subject becomes committed to a system of belief without the intervention of any rational element whatsoever. Nonetheless, I think that such a description of the situation would be confusing, and it would be better characterized as an extreme case in which we should recognise our failure to identify the subject's beliefs.

A different objection to my argument would be to claim that the interpreter does not need to take on such strong commitments, but that he need only identify what constitutes a good reason *for the subject*, regardless of whether it also constitutes a good reason *for the interpreter*. However we can only refer to the particular criteria of rationality of a subject if we have previously determined his system of belief, since this is the only way we shall be able to distinguish between what constitutes a good reason for him and what constitutes a good reason for us. Consequently in order to identify a system of belief one cannot start from what is peculiar to that system, since, *ex hypothesi*, we do not know its content at the very beginning of the interpretation. This is why in the initial stages of any process of interpretation the notions of a reason and a good reason for believing must necessarily refer to a shared rationality, common to the interpreter and to the subject who is being interpreted.

We may conclude that every interpretation of an alien system of belief must start off by assuming the basically rational character of the system in question; that is to say, by assuming that a subject will only commit himself to beliefs which are basically coherent and which, moreover, he has in general good reasons to uphold. At the beginning of the interpretation the subject must be supposed to have 'good reasons', and not only 'reasons', to commit himself to certain beliefs because the rationality, which is necessarily presupposed at this stage of the process, cannot be peculiar to the subject's system of belief but must be shared by the interpreter and the subject. This is so because it is only by making this initial assumption that

we can come to discover what is peculiar to the other system of belief, and thus appreciate certain incongruities in the system and distinguish between beliefs for which the subject has good reasons and those which he holds for bad reasons or for no reason whatsoever.

This shared rationality should not, needless to say, be understood in a rationalist sense. It should not be thought to imply the possibility of making explicit a universal standard of rationality which could be used to evaluate the rationality of any system of belief. For, in the process of interpretation, the content of such a standard of rationality can only be fixed by agreement about the rationality of certain beliefs, and not vice versa. Thus it is the agreement between the interpreter and the subject with respect to certain judgements, and not a supposed universal standard, which constitutes the core of the rationality which they must, perforce, share. Now this agreement in judgements, which is a necessary condition for all interpretation, will include, as we have seen, not only the idea of a common logic, but also that of a partially shared characterization of reality, precisely in so far as the attribution of beliefs depends on establishing what is true.

All of this enables us to see that in all interpretation there is an element shared by both the interpreter and the subject, as well as some elements which are peculiar to one or the other. These unshared elements stem from the fact that each subject shapes his particular conception of reality from his specific language, from within his own historically defined culture; in other words, he defines what is real and what is unreal within his own network of assumptions. The shared element, on the other hand, accounts for those conditions which must hold if the interpreter and the person being interpreted are to recognise each other as subjects holding beliefs; in short, as subjects.

If we now examine in slightly more detail how these two elements are related, we shall see that the process of interpretation is not merely a way of determining the contents of another subject's judgements, but is also a means of determining reality in so far as the interpreter himself may be forced to revise his own conception of reality.

II

Let us, then, consider the case of an interpreter T who seeks to interpret the system of belief SB of a subject S, to see how the process of interpretation develops. The first thing that the interpreter T has to take into account is the verbal behaviour $(VB_{.1}, VB_{.2}, \ldots VB_{.n})$ of S together with the circumstances which surround it $(C_{.1}, C_{.2}, \ldots, C_{.n})$.

Let us suppose that T begins his enquiry by interpreting $VB_{.1}$, and thus that he does not yet understand to rest of S's verbal behaviour. Here we

come up against our first problem, since, according to the previous section, a correct interpretation of VB_1 must form a part of a general interpretation which makes S's overall linguistic behaviour appear basically coherent and maximally true. Consequently we shall only be able to judge how correct the proposed interpretation of VB_1 is after working out a complete interpretation of all of S's verbal behaviour; but, obviously, we cannot reach this exhaustive interpretation unless we have previously established the meaning of each individual instance of his verbal behaviour. So the interpretation seems to be moving in a circle we cannot get out of, nor even get in to.

The only device available to T which will allow him to break into the hermeneutical circle is his own network of assumptions. Since the interpreter must assume, at first, that his own network of assumptions has some basic elements in common with that of the subject he is interpreting, and since he must interpret the subject's behaviour optimizing its rationality, then the only strategy he can follow in order to fulfil these two requirements is to interpret VB_1 in such a way that it fits in with his own system of belief. In doing this, the interpreter is using the meaning of the totality of his own linguistic behaviour to anticipate the meaning of the whole of S's linguistic behaviour.

Needless to say, as the interpreter comes across new instances of the verbal behaviour of S, he may be led to revise the meaning he initially attributed to VB_1, and so to revise his interpretation of S's global linguistic behaviour; as a result of which T might discover that certain assumptions from which he was forced to start turn out to be illegitimate in that they are distorting his interpretation. However, since the interpreter must elucidate the truth of what S says, he might well discover not only this misunderstanding, but also that some of his (the interpreter's) assumptions were false; that is to say, constituted an incorrect representation of reality. This is so because the dialogue he has set up with the subject forces the interpreter to re-examine the rationality of his own beliefs and assumptions in the light of the subject's remarks and judgements in so far as the interpretation is, perforce, committed to clarifying what constitutes a good reason for believing something about reality.

Thus, we can see that interpretation is not only a way of determining the beliefs of another subject, it is also a way of establishing what is real, what judgements are true. In this sense I think one can maintain that interpretation and understanding are structured like what may be termed an *open dialogue*: that is, a dialogue between two subjects, which, starting from partially shared networks of assumptions, is oriented towards establishing what is true, as this is the only way of determining the content of the other's judgements. In such a dialogue the possibility of revising the assumptions of either subject must of course always be left open.

This process of determining what is true will never come to an end,

since reality is always viewed from a historically defined perspective. One can always transcend one's own perspective by establishing a dialogue with other historically conditioned perspectives. In this way, the structure of understanding is more akin to a spiral than to a circle, since it never returns to its starting point but, moving from the whole to the part and from the part to the whole, is continually advancing towards an unattainable and *a priori* indeterminable end.

Be that as it may, our recognition of the historical nature of every characterization of reality or of truth should not lead us to overlook a decisive feature of the semantics of the words 'real' and 'true': when we say that such and such really is the case, or when we affirm that a statement is true, we are not only committed to its being real or true for me, for us, or for us and for you, but we are also assuming that its truth can be maintained beyond any particular epistemological perspective.

Now, given the historical nature of every characterization of reality, of what is true, how can this system-independent feature of the notions of truth and reality be incorporated into our analysis of understanding? I think this feature can be incorporated only if we consider it as if it were the result of a hypothetical idealized agreement about truth, which, although historically unattainable, serves as a guide or point of reference for all cognitive processes. On such a view the question of whether reality is interpreted as a noumenal entity present in each historical moment or as a regulative idea is irrelevant, since in both cases the notion of reality is fulfilling the same epistemological function, that of guiding knowledge; and it makes no sense, in either case, to talk of its complete characterization at any moment in history. For, although truth is not a contingent property of a statement, and although the identity of what is real is not conditioned by our historically determined knowledge, nevertheless the affirmation that a statement is true and the identification of what is real *are* historically conditioned, so that in principle any historical agreement about truth or the identity of what is real is subject to revision.

The fact that every historical characterization of what is real is carried out from within a network of assumptions which is shared by a particular community of subjects implies that every cognitive process reproduces the structure of understanding. Thus advances in scientific knowledge, in the form of causal explanations of events, cannot be thought of as being independent of the processes of understanding. The reason for this is clear if we remember that the scientific system of belief must demonstrate its rationality in a dialogue with other belief systems and must thereby accept the possibility of discovering, within this dialogical process, the irrationality of some of its assumptions. Moreover the same situation holds, though perhaps less obviously, within the scientific system of belief itself, since the content and the determination of the truth of a scientific theory cannot be thought of as independent of the agreement existing among the members

of a scientific community about what constitutes evidence for or against the theory. Even in the case in which, at the outset, we share a considerable number of assumptions, this agreement should be understood as the result of an open dialogue aimed at establishing what is real. For here too we must reckon with the possibility that someone might challenge the rationality of these initial assumptions when determining the content and truth of a scientific theory. Consequently understanding, conceived of as an open dialogue, can be regarded as a cognitive structure presupposed in every attempt to characterize what is real; and it is precisely in this sense that we can speak of the cognitive universality of understanding, as opposed to the particularity of any other feature of a given cognitive structure.

Finally, we can now see that the basic error of positivist anthropology seems to lie in its failure to realise that its commitment to a causal explanation of reality meant that its vision of reality was historically conditioned by a definite network of assumptions. This is why the positivist anthropologist did not feel it necessary to bring into the open his own assumptions or to revise his own view of reality as the result of his dialogue with other cultures.

III

The recognition of the cognitive universality of understanding and, therefore, of the fact that the scientific characterization of reality reproduces the general structure of understanding, does not necessarily rule out the possibility of offering a causal explanation of understanding itself.

One could argue, accepting all that has been said in the previous section, that at a certain moment in history the dialogical process of knowledge reaches a conception of reality according to which reality is constituted only by sequences of events which may be causally explained, so that understanding itself might also be explained causally. In this way advances in the historical dialogue to determine what is true could lead us to revise our analysis of the structure of this open dialogue.

However, the possibility of offering a causal explanation of understanding can be interpreted in various ways:

(1) The processes of understanding can be identified within a theory which employs only physical terms (excluding mental or linguistic ones). As a result cognitive processes, in so far as they can be identified in physical terms, may be subsumed under general laws and thus be causally explained.

(2) Cognitive processes can be identified only in mental or linguistic terms,

but this does not exclude *a priori* the possibility of setting up type-type correlations between classes of events identified in physical terms and classes of events identified in linguistic terms.

(3) Cognitive processes can be identified only in mental or linguistic terms, and it is not possible to establish strict nomological correlations between types of events identified in physical terms and types of events identified in linguistic terms. This, however, does not prevent us from maintaining that each particular linguistic event is identical to a particular physical event, and that different linguistic events cannot be identical to the same physical event. Thus all events are physical, although some are also linguistic. Now, as all physical events have a description which enables us to subsume them under a strict law, we can say that all events qua physical events are liable to causal explanations.

I shall now try to show that the first view is untenable; that the second, if interpreted non-trivially, is implausible; and, finally, that the third can be interpreted in such a way that it is compatible with the idea of the cognitive universality of understanding.

I think that Quine is committed to the first view when he maintains, in *Word and Object*, that it is possible to identify the meaning of a sentence by identifying the speaker's dispositions, so that a speaker's linguistic behaviour could be explained in dispositional terms.

The attribution of dispositions presupposes, for Quine, the existence of structural relations between certain physical elements, the knowledge of which would enable us to explain in nomological terms what we are currently able to explain only in terms of dispositions. As a result, the fact that dispositional sentences take the subjunctive conditional form does not make them especially unmanageable as it is possible, by definition, to establish systematic relations between their truth and the truth of certain indicative conditionals:

> Dispositions are, we see, a better-behaved lot than the general run of subjunctive conditionals; and the reason is that they are conceived as built-in, enduring structural traits. Their saving grace extends, moreover, to many subjunctive conditionals that do not happen to have acquired one-word tags like 'soluble' and 'fragile'. An example was the 'would prompt assent' of § 8. For there again a disposition was concerned, albeit unnamed: some subtle neural condition, induced by language-learning, that disposes the subject to assent to or dissent from a certain sentence in response to certain supporting stimulations.[5]

Yet, given the normative character of language, the basic assumption from which Quine starts in order to identify the meaning of sentences is, in principle, problematic. We have already seen that identifying a piece of behaviour as linguistic behaviour implies the possibility of evaluating it as

rational or irrational, as coherent or incoherent, as right or wrong; by contrast, a dispositional characterization of a physical process rules out any question of its rationality, since it is meaningless to ask about the coherence or incoherence of a physical process. Consequently a dispositional approach has to demonstrate that it is possible to identify linguistic behaviour without introducing any normative element.

One attempt to do this is to resort to the notion of 'uniformity of behaviour' as a variation on the notion 'agreement in judgements'. For Wittgenstein, agreement in judgements was at the bottom of all appraisals of the rationality of verbal behaviour, so the question of the rationality of this agreement simply could not be raised; why not then understand this agreement as empirical, thereby reducing the normativity of language to the uniformity of behaviour? In other words, why not reduce rule following to a question of behaviour in accordance with a rule?

The problem with this line of argument is that the mere identification of certain regularities of behaviour does not seem to constitute a criterion for identifying an agreement in judgements. If, in the context of a scientific conception of the world, an observer were to confine himself to describing all the regularities of nature, he would find nothing that could be termed linguistic behaviour.[6] To discover the existence of an agreement in judgements one must start from a network of assumptions which must include, apart from the idea of an object, that of a subject. For to discover an agreement in judgements is to discover an agreement in judgements on the part of certain subjects. In other words, the discovery that a subject has expressed a judgement with a certain content can only be made, as I tried to show before, by another subject who enters into an open dialogue with him about the rationality of his judgements. Thus a purely causal system of representation cannot identify the existence of linguistic communication, because to do so one must resort to notions which cannot be defined in physical terms, since they are intelligible only from the point of view of a dialogue between subjects and not from that of a subject observing an object. Accordingly, we can say that agreement in judgements is not an empirical fact in the sense that it is not a fact of nature; that is, it is not a fact that can be identified from the epistemological point of view of an external observer of objects.

Searle, who in *Intentionality* defends thesis (2), acknowledges that it is impossible to identify linguistic behaviour if one uses only terms belonging to the level of physical description. Nevertheless this does not prevent him from accepting the existence of causal relations between events identified on different levels of description, nor, therefore, from admitting the possibility of offering a causal explanation of linguistic behaviour by resorting to events identified in physical terms.

Indeed, establishing causal relations between events which have been identified on different levels of description is a straightforward, everyday

practice. Searle mentions the case of the relations established between the macroscopic and molecular descriptions of substances. Water, for example, is, macroscopically, a liquid; but, from a molecular point of view, it is a set of H_2O molecules. There is nothing about the molecular properties of water that would enable us to identify it as a liquid; if we wish to do so, we must resort to a different system of representation. Nevertheless, the fact that macroscopic and microscopic properties of a substance belong to different levels of description does not bar us from affirming the existence of causal relations between the two levels; for a causal relation obviously exists between the molecular structure of water and its liquidity. Now, if this is so in the case of water, Searle asks why we should consider more problematic the existence of causal relations between our linguistic behaviour and the physical properties of our brain:

> Singular causal statements do not entail that there is a universal causal regularity which they instantiate, but the concept of efficient causation whether intentional or not, only has applicability in a universe where a high degree of causal regularity is presupposed.[7]

In principle one could argue against this that we are justified in saying that there are causal relations between the macroscopic and molecular properties of water because nomological relations between the two types of properties are known; but we do not know of the existence of any such relations between the physical properties of the brain and linguistic behaviour.

Searle holds that such an objection is based on a false supposition, namely that in order to establish the existence of causal relations between two types of phenomena it is necessary to be able to subsume these relations under general causal laws. However for Searle it is not necessary to know these general laws, nor even to assume that they can be known: simply supposing a high degree of causal regularity to exist in the world is enough to allow us to admit the existence of causal relations.

Moreover this assumed regularity has allowed us to discover a wide range of causal regularities that link up physical phenomena with mental or linguistic phenomena. Indeed, it is absurd to deny the existence of causal regularities between certain parts of structures of the brain and certain linguistic abilities, and we can be confident that thanks to advances in neuroscience we shall increase the scope and precision of our description of such regularities. According to Searle, one cannot set *a priori* bounds on such advances in knowledge:

> And while there are no 'strict' laws, there are lots of causal regularities in the operation of Intentional causation, e.g., prior intentions cause actions, thirst causes drinking, visual experiences cause beliefs. It remains an open empirical question how these higher-level states are

realized in and caused by the operations of the brain, and an open question which of the realizations are 'type-type' and which are 'token-token'. The 'a priori' arguments that I have seen against the possibility of type, rather than of token, realizations tend to neglect a crucial point: what counts as a type is always relative to a description. The fact that we can't get type-type realizations stated in, for example, chemical terms, does not imply that we can't have type-type realizations at all.[8]

Of course I do not wish to deny the important role that our increasing knowledge in the field of neurophysiology has to play in determining the relations between linguistic behaviour and the physical world, nor do I seek to rule out the possibility of establishing certain approximate generalizations which link up these two spheres. Nevertheless none of this seems to affect the core of our argument, since the parallelism that Searle proposes comes to an end precisely at the point which most interests us: does there exist an *a priori* reason to rule out the existence of nomological causal relations between the states of our brain and linguistic behaviour, relations of the kind that exist between the macroscopic and microscopic properties of water? I think that Searle's parallelism does not answer this question. All his argument shows is that there is no reason to doubt the existence of causal relations between the brain and linguistic behaviour in so far as we presuppose a high degree of regularity in the relations existing between both type of phenomena; but he offers no evidence to show that, as happens in the case of water, we are dealing with causal relations which we can subsume under general causal laws.

I believe that it can be shown that there are *a priori* reasons for maintaining that the causal relations between the macroscopic and microscopic properties of water must be of a different kind from those which hold between the brain and linguistic behaviour, and that it would be difficult to claim that nomological causal relations may be established between the latter two. This fact, far from being trivial, shows the limits of what the study of nature can teach us about linguistic communication.

To begin with one could point out that the analogy pursued by Searle does not take into account the specific character of the criteria for identifying linguistic behaviour as opposed to the criteria for identifying the properties of water, since, although we must acknowledge with Searle that the macroscopic and microscopic properties of water belong to different levels of description, we should not lose sight of the fact that in both cases we are dealing with physical properties of objects which may be identified by an external observer. With the introduction of linguistic properties on the other hand we must, as I claimed in section II, abandon the observational point of view and adopt the internal perspective of one subject interacting with another. Thus we may say that whereas we may

observe both the molecular structure and the liquidity of water, we cannot observe linguistic behaviour externally. Consequently, if Searle's analogy is to have any force at this point, he would have to show that the specific character of the criteria for identifying linguistic behaviour is not incompatible with the possibility of establishing nomological causal relations between the brain and language. But I have not found any argument which could show this.

In thesis (3) I briefly outlined Davidson's view of the relationship between physical and mental events, which he terms 'anomalous monism'. I shall now seek to offer an interpretation of this position which makes it compatible with the thesis of the cognitive universality of understanding, although I must say that I do not know how far Davidson would go along with my interpretation.

Davidson maintains that mental events are anomalous precisely in the sense that they cannot be reduced to physical events, either in the sense of thesis (1) or in that of thesis (2). Thus it would not be possible, according to Davidson, either to define mental events in physical terms or to establish causal relations which instantiate strict general laws between events identified within a physical system of description and events identified within a mental system of description. Furthermore, the anomalism of mental events entails the anomalism of linguistic behaviour since, for Davidson, mental and linguistic events are identified by the same kind of criteria, to such an extent that a detailed identification of a subject's mental events depends on the identification of the meaning of the sentences that he utters, and vice versa.

Now, along with the anomalous character of mental events, Davidson defends a monist view with respect to the identity of events. He sharply distinguishes between the identity of events and their different possible true descriptions. From an ontological point of view, Davidson upholds the physical character of all events; but, as regards the different descriptions of events, he recognizes that the mental system of description cannot be reduced to the physical system of description. This dichotomy allows a non-exclusive characterization of the distinction between physical and mental events in the field of description. Thus, an event will be mental if and only if it has a mental description, that is, a description including verbs which express propositional attitudes; on the other hand an event will be physical if and only if it can be identified in physical terms. Accordingly it is quite licit to affirm that a given event is, at the same time, mental and physical, since the fact that it has a mental description does not rule out the possibility of its being identified in physical terms. Consequently Davidson is able to defend a non-reductionist thesis about the physical identity of mental events, since now it is possible to claim that all mental events are physical without questioning the anomalism of the events described in mental terms.

Furthermore the monist thesis seems to entail a certain cognitive universality of causal explanations, since all events, in so far as they can be identified physically, can be causally explained. But this corollary is perfectly compatible with the cognitive universality of understanding. For it does not affect the possibility of claiming that all cognitive processes, in so far as they must be identified as such within an anomalous system of description, must reproduce the general structure of understanding that I put forward in section II. Of course this general structure will be present in the network of assumptions within which causal explanations are proposed, although we should also attribute a certain universality to that more particular cognitive structure presupposed in causal explanations, in so far as we are committed to the ontological universality of causal relations.

On the other hand it is possible to accept the idea of the identity of events as being completely independent of any particular system of identification without rejecting one of the basic theses of our analysis of understanding, namely that every identification of reality is carried out from within a historically conditioned network of assumptions. This is so because the identity of events can be understood, within our analysis, as a guide to every process of determining what is real. Indeed I do not see how we could eliminate this ideal character of the notions of truth and reality in so far as we recognize that the idea of holding a statement to be true depends on the notion of evidence, and that we must always leave open the possibility of revising our decision to hold certain evidence to be conclusive evidence for a given statement.[9]

Finally, I would like to insist upon the fact that the historical awareness that we have gained in recognising that every process of identification is historically conditioned also affects ontology, understood as the discipline that endeavours to elucidate the basic categories of reality. Thus Davidson's claim that the world is made up of particular events which are causally related is a claim which, like any other, must be open to revision and, therefore, may be recognised as historically biased.

IV

Finally, I would like to go over some of the points defended in this paper.

In the first place I analysed the structure of understanding as an open dialogue between two subjects who are committed to partially shared networks of assumptions, and in which the determination of what the other says can be carried out only by determining its coherence and truth. This analysis allowed us to reject full-blooded relativism since the identification of alien belief systems, with particular criteria of rationality, is possible only if we assume the existence of a rationality that is partially shared by the interpreter and the subject to be interpreted. On the other hand we

have attempted to show that positivist anthropology failed to appreciate that its own conception of reality was historically defined. I have also maintained that the structure of an open dialogue, which is proper to understanding, is reproduced in every cognitive process; since every identification of reality depends on determining the truth of statements, and this, in turn, can only be understood as the culmination of the open dialogue which characterizes understanding. As a result of this, we spoke of the cognitive universality of understanding.

Secondly, I argued that cognitive processes cannot be reduced to causal relations inasmuch as we cannot identify them in purely physical terms and then go on to establish nomological causal relations between them and other physical events. Nor does it seem plausible to claim that nomological causal relations can be set up between events identified by physical criteria and events identified by criteria of rationality. However I have accepted the possibility of making generalizations, with an indeterminable degree of probability, to link up parts or structures of the brain with our linguistic abilities. Finally, I show that it is possible to interpret Davidson's anomalous monism so that it is compatible with my analysis of the structure of understanding and with the thesis of its cognitive universality, in spite of Davidson's commitment to the ontological universality of causal relations.[10]

NOTES

1 Evans-Pritchard (1935).

2 At this point, Evans-Pritchard (1934) refers basically to Levy-Bruhl. I wish however to explain briefly what I understand by the expression 'a basically contradictory system of thought'. We can consider the beliefs of a system of thought as being logically interrelated, and establish a gradual distinction between those beliefs which are central and those which are peripheral. The most central beliefs constitute the logical foundations of the more peripheral beliefs. The sense of the expression 'logical foundation' will vary according to the different theories of meaning. Thus a rationalist will view the central beliefs as the premises from which the more peripheral beliefs are derived, whereas a pragmatico-holistic approach will regard the more central beliefs as the basis upon which the more peripheral beliefs acquire meaning.

In the light of this distinction between central and peripheral beliefs we can say that an interpretation which presents a system of thought as basically contradictory is one which views the most central beliefs of that system as being mutually contradictory.

3 In developing the arguments which close this section and those which appear in section II, I have taken ideas principally from Wittgenstein (1960), Gadamer (1960), Habermas (1983), and Davidson (1984).

4 This approach does not preclude the possibility of showing that certain philosophical positions are self-refuting because the ideas that a philosopher upholds form a specific sub-system of beliefs within a broader system to which the philosopher is committed in non-philosophical contexts. Consequently,

claiming that a philosophical position is self-refuting does not force one to present a subject's global belief system as basically contradictory.

5 Quine (1960), pp. 223.
6 It may be interesting at this point to recall paragraph 17 of the *Monadology* of Leibniz:

> And supposing that there were a machine so constructed as to think, feel, and have perception, we could conceive of it as enlarged and yet preserving the same proportions, so that we might enter it as into a mill. And this granted, we should only find on visiting it, pieces which push against another, but never anything by which to explain perception. This must be sought for, therefore, in the simple substance and not in the composite or in the machine.

7 Searle (1983), p. 135.
8 *op. cit.*, p. 271–2.
9 At this point I think it is important to take into account the following comment of Putnam (1962):

> It is perfectly rational to make stipulations to the effect that certain statements are never to be given up, and those stipulations remain stipulations to that effect, notwithstanding the fact that under certain circumstances the stipulations *themselves* might be given up. (p. 60)

I examine this problem at length in my paper 'Are all statements revisable?' (unpublished).
10 I am very grateful to my colleagues J. L. Prades, C. Moya and V. Sanfélix for comments on the first versions of this paper. I made important changes as a result of their comments. I am also much obliged to E. Turner for his translation into English of my original Spanish version.

REFERENCES

Davidson, D. 1984. *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.
Evans-Pritchard, E. E. 1934. 'Lévy-Bruhl's Theory of Primitive Mentality', *Bulletin of Faculty of Arts*, University of Egypt.
Evans-Pritchard, E. E. 1935. 'Science and Sentiment', *Bulletin of Faculty of Arts*, University of Egypt.
Gadamer, G. H. 1960. *Wahrheit and Methode*, Tübingen.
Habermas, J. 1983. 'Rekonstruktive vs verstehende Sozialwissenschaften' in J Habermas, *Moralbewusstsein and kommunikatives Handeln*, Frankfurt: Suhrkamp.
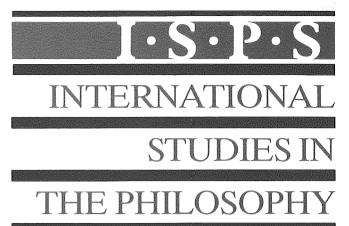**Leibniz, G. W.** 1965. *Monadology*, tr. P. and A. M. Schrecker, New York.
Putnam, H. 1962. 'The Analytic and the Synthetic', in *Philosophical Papers*, v 2, Cambridge: Cambridge University Press, 1970.
Quine, W. V. 1960. *Word and Object*, Cambridge (Mass): M I T Press.
Searle, J. 1983. *Intentionality*, Cambridge: Cambridge University Press.
Wittgenstein, L. 1960. *Philosophical Untersuchungen*, Schriften vol 1, Frankfurt: Suhrkamp.

I·S·P·S

INTERNATIONAL

STUDIES IN

THE PHILOSOPHY

OF SCIENCE

The Dubrovnik Papers

Explanation in Science