Ausonio Marras

Intentionality and Physicalism: a Resolvable Dispute

Abstract: This paper discusses the traditional antagonism between the Intentionalist and the Physicalist paradigms of the nature of mind and human behaviour. After tracing the development of the concept of intentionality in contemporary analytic philosophy and delineating some of the problems it presented for the so-called Thesis of Physicalism in several of its formulations, the paper proceeds to show how a liberalized methodology of theory construction and an understanding of functional systems and artificial intelligence models may enable us to reconstruct the intentional states of persons in a way compatible with the demands of physicalism. In particular, it is suggested that the intentional states which, on the intentionalist paradigm, mediate human behaviour might be understood as (theoretically postulated) functional states of a physical organism modelled on the functional states of a (probabilistic) automaton.

1. In contemporary analytic philosophy the concept of intentionality has played an important role in discussions on the nature of mind and human behaviour. Like many philosophical terms, the term 'intentionality' (and the cognate adjective 'intentional') has been used in more than one sense. There is a sense directly relevant to action-theory according to which 'intentionality' refers to a familiar characteristic of behaviour — its purposefulness and goal-directedness. But there is another, philosophically more technical sense, according to which intentionality is said to be a characteristic of mental functions, or of consciousness as such; it is often in terms of this latter notion that the more familiar, action-sense of 'intentionality' has often been explained, so that human behaviour is said to be intentional because it is at least partly under the control of mental functions which are, necessarily, intentional in the technical sense.

   The concept of intentionality in this latter sense has proved to be of great heuristic value as a conceptual tool for reformulating and analyzing traditional problems in philosophical psychology and action theory; but it has also provided seemingly insurmountable problems for those philosophers of scientific bent who profess a faith in the possibility of a complete and fully adequate description of man as a fully physical system. The traditional antagonism between the two paradigms of explanation of human behaviour — the 'empirical' (mechanistic, behaviouristic) model on the one hand and the 'rational' (teleological, mentalistic) model on the other — received rigorous analysis through the application of the new conceptual tools of logic and analytic philosophy, and was recast in more precise and perspicuous form as a debate between Physicalism and Intentionalism. This paper will attempt to trace, in a summary and highly selective way, the development of the concept of intentionality in contemporary analytic philosophy, delineating some of the problems it presented for the so-called *Thesis of Physicalism* in some of its

different formulations, and to advance some reasonable conjectures about the pros-
pects for a resolution of the Intentionalist-Physicalist controversy through a physi-
calistic 'reconstruction' of the concept of intentionality.

2. The philosophical concept of intentionality goes back to the Scholastic doc-
trine of the *esse intentionale* which a known object is said to have in the knower.
According to Thomas Aquinas, this 'intentional presence' comes about through a
modification of the subject (*immutatio*) who, in the knowing act, takes on some-
how the form or *species* of the object. The species, by 'inhabiting' the subject not
only as an accidental modification of it (*secundum esse naturale*) but also, essen-
tially, as a representation of the object (*secundum esse repraesentativum*), becomes
the vehicle of the 'intentional reference' to the object which, as Brentano was later
to point out, constitutes the very essence of consciousness.[1]

It is precisely through the work of Brentano that analytic philosophy rediscover-
ed the importance of the concept of intentionality. This was no accident. For it
must be remembered that, after the second world war, a critical re-examination of
certain 'dogmas' of logical empiricism was taking place in several Anglo-American
philosophical circles. One of these dogmas was the *thesis of the unity of science.*
According to one of Karl Hempel's formulations, this thesis asserted that "all the
branches of science are in principle of one and the same nature; they are branches
of the unitary science, physics" (Hempel 1935).[2] Now, as is well known, in his
Psychologie vom empirischen Standpunkt (1874), Brentano had asserted, in effect,
the autonomy of psychology. More precisely, he had asserted that psychical pheno-
mena, *in as much as they are intentional*, that is, in as much as they 'bear a refe-
rence' to an object, are irreducible to physical phenomena. Thus, for Brentano,
intentionality constitutes a valid criterion for distinguishing psychological from
physical phenomena, and thus for distinguishing psychology from physics. Unsur-
prisingly, therefore, the Brentanian criterion was taken up by certain philosophers
in the hope that it might turn out to be a valid instrument for a critique of the
aforementioned empiricist thesis.

But let us look at the matter more closely. Let us keep in mind, first of all, the
linguistic orientation of logical empiricism and analytic philosophy in general —
that is, their tendency to view the philosophical enterprise as an investigation or
analysis of various linguistic forms. In accordance with this approach, the thesis of
the unity of science received this linguistic formulation by Rudolf Carnap in 1935:
"[The] physical language is the basic language of all science, . . . a universal language
comprehending the content of all other scientific languages. In other words, every
sentence of any branch of scientific language is equipollent to some sentence of the
physical language, and can therefore be translated into physical language without
changing its content." (89) The behaviouristic movement, which endeavoured
to analyse psychological concepts in terms pertaining to observable behaviour
(e.g. 'thinking' in terms of 'subvocal speaking', or in terms of dispositions to make
certain utterances), was viewed by many empiricists as a concrete, if only partial,
expression of their physicalistic aspirations. Moreover, since the syntax of a physi-
calistic language was supposed to mirror the logical structure of a symbolic language

such as that of *Principia Mathematica*, which is a strictly *extensional* language, it was assumed by many (for example by Carnap) that the physicalistic language ought to be, or at least be reducible to, a strictly extensional language. The thesis of the unity of science, that is to say, of physicalism, thus also carried with it a *thesis of the extensionality of the language of science*. This is how Carnap formulated it in 1937: "For every given language S_1 an extensional language S_2 may be constructed, such that S_1 may be translated into S_2." (320)

In order to conduct an effective critique of such a linguistic thesis, it was obviously expedient to carry on the discussion at a linguistic level. And this is precisely what the critics of physicalism did. At first they objected, not unreasonably, that the proposed physicalistic translations of psychological statements were simply unacceptable (*critique of behaviourism*); later, however, their attack took a more radical and effective form when they came to formulate a positive thesis of their own on the nature of psychological language which, if valid, would have demonstrated the *impossibility in principle* of the physicalistic thesis. It is clear that one could not refute the physicalistic thesis by merely pointing out the inadequacy of specific physicalistic translations; for the physicalist could always reply that his translations were merely *proposals*, translation *sketches*, which, with the improvement of methodology and experimental techniques in the behavioral and neurophysiological sciences and the acquisition of new data, laws and theories, would become more precise, more complete and therefore more satisfactory. Thus the need was felt, among the critics of physicalism, for a more positive and radical approach to the problem.

3. Roderick M. Chisholm, of Brown University, was probably the first to advance a thesis whose validity would have demonstrated the impossibility of the physicalistic thesis (as previously formulated). His thesis made use of the Brentanean notion of intentionality which, however, was re-interpreted in linguistic terms. Recall that according to Brentano what distinguishes physical from psychical phenomena is the intentionality of the latter, and that every psychical phenomena, *qua* intentional, implies "a reference to a content, a direction upon an object (by which we are not to understand a reality in this case) . . ." (Brentano 1874, 124). In other words, not only must every psychical phenomenon make reference to an object, but such an object may not even exist in reality. One may think of Pegasus, which does not exist, but one cannot ride on it. One may be afraid of a ghost, which does not exist, but one cannot capture it.

In the light of this interpretation of the Brentanean conception of intentionality, this is how Chisholm went about formulating his own linguistic thesis about the concepts of psychology.[3] Instead of taking phenomena as the object of his investigation, he takes certain linguistic forms, certain types of sentences, of which, on the basis of certain criteria presently to be discussed, he distinguishes two irreducible categories: *intentional sentences and non-intentional sentences*. (Note that the attribute 'intentional' in Chisholm's analysis applies properly to linguistic items.) Then Chisholm formulates the following thesis: It is not necessary to use an intentional language (that is, a set of intentional sentences) in order to describe non-psychical,

or physical, phenomena; on the other hand, in order to describe psychical phenomena (such as thinking, believing, perceiving, knowing, wanting, hoping, etc.), it is necessary to use an intentional language (or at least a 'technical' language that can only be explicated in terms of an intentional language) (Chisholm 1957; 1958, ch. 11). Chisholm's distinction between an intentional and a non-intentional language is meant to reflect the logical empiricist distinction between the language of psychology and the physicalistic language; moreover, if valid, the distinction would demonstrate the *untranslatability* of the one language into the other.

How, then, does Chisholm distinguish between intentional and non-intentional sentences? According to a definition given in 1957, he defines as intentional those sentences, and *only* those sentences, which satisfy at least one of the following criteria (each of which is intended as a sufficient condition) (Chisholm 1957; 1958, ch. 11).

(1) A sentence is intentional if it contains a substantival expression such that neither the sentence nor its negation implies that the substantival expression refers to something existent or to something non-existent. For example, the sentence 'Diogenes sought an honest man' is intentional because neither this sentence nor its negation ('Diogenes did not seek an honest man') imply either that there is or that there is not an honest man. On the contrary, the sentence 'Diogenes lit his lantern' is not intentional by this criterion because the sentence *does* imply the existence of a lantern.

(2) A sentence is intentional if it contains a subordinate clause such that neither the sentence nor its negation imply either that the subordinate clause is true or that it is false. For example, the sentence 'Caesar believed that Brutus would be faithful' is intentional because neither this sentence nor its negation imply that Brutus would be faithful nor that he would be unfaithful. On the contrary, the sentence 'Brutus's unfaithfulness prevented Caesar from becoming emperor' implies that Caesar did not become emperor, and thus the sentence is not intentional by this criterion.

(3) A sentence is intentional if it contains a substantival expression such that, by replacing this expression with another equivalent expression,[4] it is possible to alter the truth value of the sentence. For example, the sentence 'Peter knows that Caesar wrote *De Bello Gallico*' may be true while the sentence 'Peter knows that Calpurnia's husband wrote *De Bello Gallico*' may be false, despite the identity of Caesar with Calpurnia's husband. Thus the sentence is intentional. On the contrary, the sentence 'Brutus caused Caesar to be murdered' is not intentional by this criterion because its truth value is not altered if we substitute the expression 'Calpurnia's husband' for the expression 'Caesar'.

4. Chisholm's thesis caused a great deal of controversy.[5] On the one hand there was a question as to whether the class of intentional sentences as defined by Chisholm's criteria actually coincided with the class of psychological sentences. In other words, there was a question as to whether Chisholm's criteria of intentionality fulfilled the end for which they were designed, which was to show that *all* sentences of psychology, and only such sentences, have the property of intentionality, and thus show that the sentences of psychology may be unequivocally distinguished

from the physicalistic sentences. It was pointed out, for example, that modal sentences such as 'It is logically possible that the morning star is not identical with the planet Venus' turn out to be intentional by the third criterion (as one can see if we replace in our example the expression 'the morning star' for the equivalent expression 'the planet Venus'); and yet, surely, modal sentences do not belong to psychology. On the other hand renouncing the third criterion would have the unhappy consequence that one could not demonstrate the intentionality of 'cognitive' sentences (sentences containing such verbs as 'know', 'observe', 'realize', 'demonstrate', 'perceive', 'remember', etc.): such sentences, though surely psychological, could not be shown to be intentional by the first two criteria.

In the face of these and other objections, Chisholm was forced to seek new criteria of intentionality (Chisholm 1963-4; 1967). These, however, turned out to be no less problematic than the former (see Luce 1964-5; Sleigh 1964-5; Lycan 1969); moreover, in the opinion of the present writer, the new criteria were further removed from Brentano's helpful intuitions on intentionality, which were instead embodied, in linguistic guise, in Chisholm's first two criteria. Precisely because of this, still in the opinion of the present writer, it is preferable to accept the first two criteria which, as has been argued elsewhere (Marras 1968; 1980),[6] may well turn out to be adequate to the task if strengthened by the following condition: 'A sentence is intentional if it entails another sentence that satisfies either the first or the second criterion.' In this way it is possible to demonstrate the intentionality of cognitive sentences (e.g., of the sentence 'Peter knows that Caesar wrote *De Bello Gallico*', which entails the intentional sentence 'Peter *believes* that Caesar wrote *De Bello Gallico*'), without having to make use of the problematic third criterion, which, as we shall see, constitutes not so much a sufficient condition of *intentionality* as rather a sufficient condition of *non-extensionality* (or intensionality).

5. There is no need, however, to dwell further on these controversies. Let it simply be remarked that Chisholm's program to formulate precise criteria of intentionality was generally thought to be arduous, but in principle realizable; and many, more or less ingeniously, endeavoured to bring the task to completion.[7]

On the other hand, even admitting the realizability of Chisholm's program, there were many who questioned its utility (see, e.g., Nochlin 1953 and Cornman 1966). It will be appropriate, therefore, to re-examine the *point* of Chisholm's thesis, at least in the perspective of the physicalistic program as described thus far.

We can at least assert the following. If we could show that there are precise logical criteria of intentionality which all and only psychological sentences are able to satisfy, then we would have shown that psychological sentences constitute a logically distinct and irreducible class of sentences, so that any attempt to *translate* psychological sentences into physicalistic (and thus non-intentional) sentences would be bound to fail. As a consequence Chisholm's thesis would, if valid, entail the impossibility in principle of the physicalistic thesis in the formulation thus far considered. Moreover, since it is a *necessary* condition of a sentence's *extensionality* that its truth value should not be altered as a result of substituting an expression therein contained with another equivalent expression, it follows that Chisholm's third cri-

terion of intentionality, which is the equivalent of the *negation* of the extensionality condition just mentioned, constitutes a *sufficient* condition of *non-extensionality*; and since many psychological sentences satisfy Chisholm's criterion, they are to be classified as non-extensional (or intensional). Chisholm's thesis, therefore, also entails the falsity of the thesis of the extensionality of scientific language; at least some of the sentences of psychology are non-extensional.

While recognizing the soundness of the foregoing remarks, one may still raise the following two objections which, if valid, call into question the very usefulness of the Chisholmian program — or of *any* intentionalist program which, like Chisholm's, is concerned with establishing the logical distinctiveness of the psychological language.

The first objection is as follows. We have been restricting our use of the terms 'psychological concepts' and 'psychological sentences' to refer to the concepts and sentences of a common-sense *mentalistic* psychology, insofar as these apply to 'states of mind' such as thinking, doubting, wishing, intending, expecting, and the like. But, the objection goes, a behaviouristic psychology as an *empirical* science has no *use* for such concepts, for its proper task is the *explanation* of *overt* human behaviour, not the *analysis* of outmoded *mentalistic* concepts. Since — so behaviourists claim — human behaviour can (and must) be explained without the use of mentalistic concepts, the Chisholmian thesis that psychological (mentalistic) sentences are irreducibly intentional should constitute no threat to scientific behaviourism (physicalism).

The reply to this objection has three parts. (1) If Chisholm's definition of intentionality is revised in the way previously suggested (p. 5), then the class of intentional and (thus) psychological sentences will include those that refer to purposeful, deliberate human action: for the sentence 'John raised his hand' as distinct from 'John's hand went up' normally entails that John *wanted* or *intended* to raise his hand, or *had a reason* for raising it, etc. But if 'John raised his hand' is *both* intentional *and* describes a phenomenon within the proper domain of a behaviouristic psychology (namely an overt human action), then this fact *does* constitute a threat to behaviourism, in so far as his subscribes to the physicalistic thesis as previously characterized.

(2) Even if the proper task of a scientific psychology were the explanation of *overt* human behaviour, it has become increasingly clear — even among behaviourists (see, e.g., Dulany 1968) — that such behaviour cannot be explained without the use of (explicitly or implicitly) mentalistic mediating constructs. Facts pertaining to perceptual organization, learning sets, memory, attention, awareness of rules, motivation, etc. have disrupted the former simplicity of S-R connections: responses are conditioned not simply to physically definable stimuli but to *perceived* (perceptually organized) stimuli; S-R connections are reinfocred not simply by physically definable reinforcers but by *desired* or *valued* reinforcers; and so on. Whether these facts can be *fully redescribed* in a purely physicalistic, extensional language is precisely what someone like Chisholm would deny.

(3) Whether or not such mental phenomena as thinking, believing, desiring and the like play a role in the explanation of behaviour, no one — not even a behaviourist — would deny that such phenomena *exist*; and if they exist they fall within the scope of scientific investigation, even if they are outside the scope of a behaviouristic psychology. Hence if physicalism (as previously characterized) were true, it should be possible to describe such phenomena in a physicalistic, extensional language — and this, again, is something that someone like Chisholm would deny.

If the above objection fails, there is a more powerful objection that can be raised against the Chisholmian program — an objection which results from reflecting on the historical, evolving character of the physicalist thesis. For it must be remembered that in the last few decades the neo-empiricist movement underwent a period of self-examination and self-criticism which has led to a liberalization of its more rigid earlier positions. The thesis of the extensionality of scientific language, for example, is no longer universally accepted by the neo-empiricists; instead, many of them now generally recognize that sentences such as those which express universal laws or which attribute dispositional properties (solubility, inflammability, etc.) presuppose modal concepts such as physical necessity or causal implication, and are thus likely to have an irreducibly non-extensional logical structure. The non-extensionality of psychological sentences, therefore, should no longer be regarded as an inadmissible hypothesis by the neo-empiricists.

The thesis of physicalism has also been considerably liberalized. As early as in 1936—37, in his classic essay *Testability and Meaning,* Rudolf Carnap gave up the claim that all scientific statements should be *translatable* into physicalistic statements; instead, he advanced the weaker thesis that all scientific concepts, including those belonging to psychology, should be *reducible* to the concepts of a physicalistic language (now simply understood as the 'the thing-language', the language, that is, "which we use in everyday life in speaking about the perceptible things surrounding us" (Carnap 1936—7, 466). Whereas the requirement of translatability presupposes the complete and explicit *definability* of all scientific concepts in terms of the physicalistic language, the requirement of reducibility only involves the possibility of specifying, in a limited and provisional way (as befits scientific hypotheses) some *empirical criteria of application* for scientific terms on the basis of experimentally discovered regularities obtaining among phenomena. One way in which these empirical criteria can be specified is by means of what Carnap has called 'reduction statements', that is, statements having the following general form: 'When such and such observable test conditions are realized, an organism is in such and such psychological state if and only if such and such behavioural response or physical state is present.' Since reduction statements do not in general have the form of an equivalence '... $\equiv$ _ _ _ ' but only the form of a (complex) conditional '$***$ → (... $\equiv$ _ _ _)', they do not purport to give a complete definition of psychological terms but only partial criteria of application for the terms in question, relative to given experimental situations. Clearly, such partial empirical criteria do not exhaust the meaning of psychological terms but merely serve to 'introduce' them into the language of science on the basis of determinate experimental situ-

ations, while at the same time leaving open the possibility of further specifications of their meaning on the basis of different experimental situations.

Now the thesis of physicalism understood as a reducibility thesis as opposed to a translatability thesis does not seem inconsistent with Chisholm's thesis about the intentionality of the language of psychology. A psychological sentence may well satisfy Chisholm's criteria of intentionality and at the same time remain anchored to a solid physicalistic basis.

Further developments in empiricist thinking on the structure of scientific theories and the status of psychological concepts have further loosened the ties between such concepts and their observational physicalistic basis. In "The Methodological Character of Theoretical Concepts" (1956) Carnap came to realize the inadequacy of the earlier, strictly *dispositional*, account of psychological terms and recommended that they be interpreted instead as *theoretical* terms. The attractiveness of this view becomes obvious in the light of the more liberal conception of theoretical terms, current since Carnap's 1956 paper, according to which such terms are introducible not simply by explicit, operational definitions in terms of observables, nor even simply by means of reduction sentences, but rather by means of (i) a set of theoretical postulates which 'implicitly define' their function in the theory, and (ii) a set of 'correspondence rules' (C-rules) which nomologically relate *some* (not necessarily *all*) of the theoretical terms to the physicalistic observation basis (see, especially, Carnap 1956; also Hempel 1965). These rules need not even take the form of universal laws; indeed, the more general form would be, for Carnap, that of statistical-probabilistic laws — a form eminently suited for the laws of psychology. The crucial point, as far as the thesis of physicalism is concerned, is that it no longer requires that all scientifically significant terms be reducible to an observational physicalistic basis; a term may be scientifically significant if it can be construed as a theoretical term in the sense just indicated. Psychological terms so construed need no longer be taken to refer to behavioural dispositions: rather, given a realistic interpretation of theoretical terms, they may be taken to refer to non-observational, *postulated* states of organisms nomologically related to the organisms' behavioural states and dispositions — the latter merely providing *probabilistic* evidence for the postulated states. Clearly, nothing prevents psychological terms thus construed from occurring in sentences which satisfy Chisholm's intentionality criteria: a psychological sentence may well be intentional and yet meet the demands of a liberalized physicalism.

6. The thesis of physicalism as now understood leaves it entirely open whether the organism's theoretical states described by the terms of a psychological theory are micro-structural states of the central nervous system or, perhaps, functional states of the whole organism, or a combination of both. Indeed, as a methodological thesis concerned with general principles of concept formation, the thesis of physicalism does not dictate what particular form a psychological theory must take. Even scientific behaviourism, once freed of its early positivistic constraints, may provide a viable theoretical framework for a physicalistic reconstruction of psychological, intentional concepts. Behaviourism, if it is to remain an *empirical*

science and thus satisfy the requirements of physicalism, must indeed require that the *set* of all its theoretical concepts be 'anchored' to a basic observational vocabulary — in particular, to a physicalistic vocabulary pertaining to overt behaviour. However, it need not require that *each* of its concepts be so anchored, nor that the anchoring be by means of explicit (operational) definitions or even by means of reduction sentences; it merely suffices that *some* of its concepts be linked to the observation basis by means of probabilistic C-rules so as to guarantee that the theory have observational consequences. In contrast with 'single-stage' ('empty-box') behaviourism — either of the classical Watsonian variety or of the operant Skinnerian variety — theoretically richer forms of 'multi-stage' (neo) behaviourism have recently been proposed (see, e.g., Osgood 1968)[8] which introduce mediating structures and mechanisms as theoretical constructs, thus providing a richer framework for explaining complex human behaviour, in the context of which the concept of intentionality might be reconstructed compatibly with the requirements of physicalism.[9]

In the spirit of a liberalized methodology of theory construction, Wilfrid Sellars has proposed a reconstruction of psychological concepts as theoretical concepts, explicating the intentionality of mental functions on the model of linguistic meaning — the latter, in turn, being interpreted by reference to the functions that linguistic expressions play in the *behaviour* of a linguistic community.[10] Although in the opinion of the present writer Sellars' account remains unsatisfactory to the extent that it is grounded on a still too impoverished S-R conditioning model of language learning and linguistic behaviour, there is no reason in principle why the latter account could not be enriched in accordance with the more liberal principles of theoretical (neo) behaviourism.[11] Having made use of the rich methodological resources of theory construction for the introduction of the mentalistic framework, there is no reason why Sellars should not make use of the same methodological approach for the introduction of theoretical concepts referring to whatever underlying cognitive structures and mechanisms might prove desirable — in view of current psycholinguistic research — for the explanation of linguistic behaviour. Only if we start with a rich enough account of linguistic behaviour can we hope to reconstruct in terms of it the intentional characteristics of mental functions.

D. M. Armstrong, a 'central state' materialist who identifies the functions of the mind with certain chemico-physical functions of the brain, also recognizes that mental functions are typically intentional, and that their intentionality is not to be eliminated in the process of analysis along with the 'residue' of the analysis, but requires instead an explanation, a theoretical reconstruction (Armstrong 1968). Armstrong's own physicalist reconstruction, following the inspiration of cybernetic models, is essentially in terms of the theory of self-regulating systems: the intentionality of a psychological state is nothing but a chemico-physical function of the central nervous system apt to produce various forms of goal-directed behaviour which may be modified and redirected by the 'experiences' (negative feedback) of the individual organism in response to the contingencies of the environment. Intentional states are thus for Armstrong inner *causal* states of an organism: they de-

termine the organisms' responses, given appropriate stimulus conditions, and thus *mediate* the manifestation of goal-directed dispositions. While not behaviouristic in the classical sense, Armstrong's views can probably be accommodated within the framework of a liberalized form of theoretical (neo) behaviourism.

7. Although, according to the physicalist, the intentional states of an organism can only be *realized* or *tokened* by physical states of a physical system, it is possible to identify and describe intentional state *types* in abstraction from any particular physical realization. I should like to conclude this paper by considering — all too briefly and somewhat generically — a currently influential physicalistic approach to the mental generally referred to as 'functionalism'.[12]

In its more sophisticated forms, this approach takes its inspiration from automaton-theory and attempts to construct a computational model of the mind. The intentional states of an organism are, on this proposal, analogous to the logical states of a (probabilistic) Turing Machine ('probabilistic automaton'). Although, strictly speaking, a Turing machine is an abstract mathematical object consisting of an ordered set on n-tuples of symbols, it is useful to think of these symbols as representing the inputs, outputs, and discrete internal states of a computing machine. For every such machine there corresponds a Machine-Table (program) which specifies, for every possible combination of inputs and internal states, the respective probabilities of the machine's outputs and transitions to the next state. The Machine-Table describes the *functional organization* of the machine, whose states are understood as functional states 'implicitly' specified by the transition probabilities given by the Machine-Table (cf. Putnam 1967).

Now, in conceiving of an organism as a physical realization of a Turing machine, the analogy suggested by the functionalist is that inputs be thought of as stimulus conditions, outputs as behavioural responses, and machine states as complex intentional or psychological states. Then, just as the machine's present state and input determine (the probability of) the next state and output, so similarly the present intentional state and stimulus conditions of an organism determine its next intentional state and behavioural response. In particular, the intentional state-types of an organism can be wholly *individuated* in terms of their systematic functional relations to certain inputs, outputs and other internal states, relative to the organism's functional organization. The intentional states of an organism are thus essentially like the machine-states of a system relative to some (empirically adequate) functional description of the system. And since the internal states of a system are characterized exclusively in terms of their extrinsic functional relations to other states, inputs and outputs, the specific *intrinsic* nature of these states in any physical realization of the system is left (largely) open: it is up to neurophysiology, biochemistry, etc. to tell us what particular structures in the organism are able to causally discharge the relevant psychological (intentional) functions.

Clearly, this account of intentional states is entirely compatible with the requirements of physicalism. The fact that intentional states can be characterized in *abstraction* from any particular physical realization of them does not deny that such states are, as a matter of fact, complex *functions* of *physical* states. A functional

('intentional') description of a system may be *neutral* with respect to any physical (structural) description of the system without being thereby *incompatible* with it. Physicalism need not deny the distinction between a functional (intentional) description of a system and a physical description of the functional *components* of the system (cf. Dennet 1971). Moreover, since a functional description is ultimately a description of how various functional components of a system *interact* with one another and with other objects and conditions in the environment, a functional description easily lends itself to a *causal* interpretation: the states it describes are specified in part in terms of input-output relations, which, as we have seen, can be interpreted as, e.g., stimuli and responses.[13] Thus, on the reasonable assumption that causal roles can only be discharged by *physical* structures, a functionalist account of intentional states implies, on some plausible *realist* (non-instrumentalist) interpretation, some form of physicalism.

A functionalist account of intentional states is also compatible with Chisholm's thesis of intentionality. For to talk of psychological states in terms of their intentional characteristics is, as Brentano and Chisholm taught us, to talk about them in terms of their *relations* to other states and other (actual or possible) objects; but functional descriptions of a machine's internal states display in general similar structural characteristics, so that the latter states can be said to 'mimic' the former. Thus just as a person may be said to *think that p* without thereby *thinking that q* even when 'p' and 'q' have the same truth value, so similarly a machine may be in a state such that, given a certain input, e.g. the punching in of 'Who wrote *Huckleberry Finn*?', it may respond with (print out) 'Mark Twain' but not 'S. L. Clemens' even though in fact 'Mark Twain wrote *Huckleberry Finn*' and 'S. L. Clemens wrote *Huckleberry Finn*' have the same truth value. For the information 'Mark Twain = S. L. Clemens' may not be stored in the machine, so that even if the Machine-Table allows for substitutivity of identicals, the substitution could not on this occasion be executed. Again: the intentionality of a perceptual state — such as perceptually *taking* something to be a sheep when no sheep is there to be perceived — might be 'mimicked' by a 'perceiving machine' which prints out 'sheep' when instructed to print out the name of the object in the center of its visual field, when in fact such an object is a bush and not a sheep: the pattern-recognition mechanism built into the machine may not be sufficiently fine-grained to discriminate sheep from bushes at certain distances.[14]

8. It is now time to sum up. I have argued that the intentionality of mental states which Brentano and the Scholastics insisted on and which Chisholm attempted to capture in linguistic form need not and must not be denied by contemporary physicalism. Chisholm's work, in particular, has forced analytic philosophers to pay close attention to the concept of intentionality, whose importance and heuristic value for the analysis of psychological concepts is now generally recognized even by philosophers who have remained faithful to the physicalist empiricist tradition. For many of these philosophers the crucial question is no longer whether psychological phenomena do or do not possess the characteristic of intentionality; granted that they possess it, the problem is to provide an *explana-*

tion of this fact in physicalistic terms. The concept of intentionality, in other words, is not to be *analyzed away*: any physicalist philosopher worth his salt must instead strive to provide a reconstruction of it in the context of a physicalistic account of man and his place in the world. Once freed of its early positivistic constraints, contemporary physicalism can strive to reconstruct the intentionality of mind — perhaps along the lines previously illustrated — as a functional, organizational feature of organisms as complex physical systems. The systematic study of the application of cybernetic models, of artificial intelligence, of communication and information theory, psycholinguistics and cognitive psychology etc., has provided a broad and interdisciplinary research area in which contemporary physicalism can hope to find a theoretical basis for a scientific account of man as a conscious, intentional, and yet fully physical organism. The Scholastic and Brentanian insight on intentionality need not be lost.

Notes

*   This research was supported by a grant from the Canada Council.

1 For a detailed account of the relation between the Scholastics' and Brentano's conception of intentionality, see Marras (1976).

2 Because physics was regarded as the basic science, the thesis was also generally referred to as *the thesis of physicalism*.

3 The first formulation dates back to Chisholm (1952). More precise and detailed formulations appeared later in Chisholm (1957), (1958).

4 Two substantival expressions are equivalent when they denote the same object.

5 See the studies by Farrell (1954–55), O'Connor (1967), Cornman (1962), (1964), Brown (1963, 1964–5), Clark (1965), Heidelberger (1965–6), Marras (1968), Rosenthal (1973).

6 Marras (1968), (1980). Chisholm now acknowledges the soundness of my proposal (private communication).

7 See, for example, Lycan (1969), Marras (1968), Sanford (1970), Morick (1976). See also, in the context of somewhat different philosophical concerns, Anscombe (1968).

8 Various kinds of behaviouristic theories are formally distinguished in Tuomela (1973); the author also argues for the indispensibility of theoretical concepts for the purpose of inductive systematization and explanation.

9 This kind of liberalization of behaviouristic methodology may help to politically 'de-sensitize' the behaviouristic program. The conditioning and control of human behaviour by the mere manipulation of environmental variables is no longer seen as a realizable task, for it is no longer the environment which plays the crucial role in learning, but the subject's own *representation* of the environment; and that representation is essentially a mediating *cognitive* structure which constitutes the subject's personal and active contribution to the learning ('conditioning') process. The 'freedom and dignity' of the individual need not be challenged by the neo-behaviourist.

10 A classic and concise statement of this thesis is contained in Sellars (1956).

11 For further discussion on this see Marras (1973a), (1973b), (1973c), (1977), (1978); Sellars (1973); and Tuomela (1977), chapter 3, section 5.

12 See, e.g., Putnam (1967), Fodor (1968), Lewis (1969), Dennet (1971). For a survey of various forms of functionalism see Elugardo (1979). An improved Sellarsean version of functionalism is given by Tuomela (1977). The type of functionalism described below may be referred to as '(Turing) Machine Functionalism'.

13 The old saw, once current among intentionalists (e.g. Melden) that human action is to be explained in terms of reasons *and that reasons can't be causes*, has long been discredited: the *acceptance* of something as a reason, the *belief* that such and such is conducive to certain goals, can certainly be a cause of action. Such states as beliefs and desires are theoretically

postulated as causal ingredients of action precisely in the endeavour to explain how organisms can discharge those intentional functions describable in functionalist terms.

14 Think of the perceiving machine as a computing device programmed to perform a many-to-one mapping of stimuli-patterns onto sets of appropriate responses (print-outs). The mapping is mediated by a pattern-recognition mechanism which, in its crudest form, might consist of a system for matching stimulation-patterns projected onto a mosaic of sensitive cells with patterns stored in the machine's 'memory' (e.g. in a template). A variety of factors (distorted projections, interferences, weakening of the cells, defective template, etc.) could cause the system to malfunction, and thus to 'misperceive'.

Bibliography

Anscombe, E. (1968), The Intentionality of Sensation: a Grammatical Feature, in: Butler, R. J., ed., *Analytical Philosophy*, 2nd Series, Oxford

Armstrong, D. M. (1968), *A Materialist Theory of the Mind*, London

Brentano, F. (1874), *Psychologie vom Empirischen Standpunkt*, Leipzig

Brown, S. C. (1963), Intentionality Intensified, *Philosophical Quarterly* 13, 357—360

— (1964—5), Intentionality without Grammar, *Aristotelian Society Proceedings* 65, 123—146

Carnap, R. (1935), *Philosophy and Logical Syntax*, London

— (1936—7), Testability and Meaning, *Philosophy of Science* 3—4, 419—471, 1—40

— (1937), *The Logical Syntax of Language*, New York

— (1956), The Methodological Character of Theoretical Concepts, in: Feigl, H. and Scriven, M., eds., *Minnesota Studies in the Philosophy of Science I*, Minneapolis

Chisholm, R. M. (1952), Intentionality and the Theory of Signs, *Philosophical Studies* 3, 56 —63

— (1957), *Perceiving: A Philosophical Study*, Ithaca, N. Y.

— (1958), Sentences about Believing; reprinted in Marras (1972)

— (1963—4), Notes on the Logic of Believing; reprinted in Marras (1972)

— (1967), Brentano on Descriptive Psychology and the Psychological, in: Lee, E. N. and Mandelbaum, M., eds., *Phenomenology and Existentialism*, Baltimore

Clark, M. (1965), Intentional Objects, *Analysis, Supplementary Volume* 25, 123—128

Cornman, J. W. (1962), Intentionality and Intensionality; reprinted in Marras (1972)

— (1964), The Extent of Intentionality, *Philosophical Quarterly* 14, 355—357

— (1966), *Metaphysics, Reference and Language*, New Haven

Dennett, D.C. (1971), Intentional Systems, *Journal of Philosophy* 68, 87—106

Dulany, D. (1968), Awareness, Rules, and Propositional Control: A Confrontation with S-R Behavior Theory, in: Dixon, T. and Horton, D., eds., *Verbal Behavior and General Behavior Theory*, Englewood Cliffs, N. J.

Elugardo, R. (1979), *Functional Materialism and the Qualia Objection*, doctoral dissertation, University of Western Ontario

Farrell, B.A. (1954—5), Intentionality and the Theory of Signs, *Philosophy and Phenomenological Research* 15, 500—511

Fodor, J.A. (1968), *Psychological Explanation*, New York

Heidelberger, H. (1965—6), On Characterizing the Psychological, *Philosophy and Phenomenological Research* 26, 529—536

Hempel, K. (1935), Logical Analysis of Psychology; reprinted with revisions in Marras (1972)

— (1965), *Aspects of Scientific Explanation*, New York

Lewis, D. (1969), An Argument for the Identity Thesis, in: Rosenthal, D. M., ed., *Materialism and the Mind-Body Problem*, Englewood Cliffs, N. J., 1971

Luce, D.R. (1964—5), On the Logic of Belief; reprinted in Marras (1972)

Lycan, G. (1969), On 'Intentionality' and the Psychological; reprinted with revisions in Marras (1972)

Marras, A. (1968), Intentionality and Cognitive Sentences; reprinted with revisions in Marras (1972)

— ed. (1972), *Intentionality, Mind and Language*, Urbana, Ill.

— (1973a), Sellars on Thought and Language, *Noûs* 7, 152—163

— (1973b), On Sellars' Linguistic Theory of Conceptual Activity, *The Canadian Journal of Philosophy* 2, 471—483

— (1973c), Reply to Sellars, *The Canadian Journal of Philosophy* 2, 495—501

— (1976), Scholastic Roots of Brentano's Conception of Intentionality, in: McAlister, Linda L., ed., *The Philosophy of Brentano*, London

— (1977), The Behaviourist Foundation of Sellars' Semantics, *Dialogue* 16, 661—675

— (1978), Rules, Meaning and Behaviour: Reflection on Sellars' Philosophy of Language, in: Pitt, J., ed., *The Philosophy of Wilfrid Sellars: Queries and Extensions*, Dordrecht, Holland

— (1980), Intentionality Revisited, *Philosophia*, forthcoming

Morick, H. (1976), Extensionalizing the Non-Psychological, *Philosophy and Phenomenological Research* 36, 551—553

Nochlin, P. (1953), Reducibility and Intentional Words, *Journal of Philosophy* 50, 625—638

O'Connor, D.J. (1967), Tests for Intentionality, *American Philosophical Quarterly* 4, 173—178

Osgood, C. (1968), Toward a Wedding of Insufficiencies, in: Dixon, R. and Horton, D., eds., *Verbal Behavior and General Behavior Theory*, Englewood Cliffs, N.J.

Putnam, H. (1967), The Nature of Mental States, in: Rosenthal, D. M., ed., *Materialism and the Mind-Body Problem*, Englewood Cliffs, N. J., 1971

Rosenthal, D.M. (1973), Talking about Thinking, *Philosophical Studies* 24, 283—313

Sanford, D.H. (1970), On Defining Intentionality, *Proceedings of the XIVth International Congress of Philosophy*, vol. 2, Vienna

Sellars, W. (1956), Empiricism and the Philosophy of Mind; reprinted in: Sellars, W., *Science, Perception and Reality*, London 1963

— (1973), Reply to Marras, *Canadian Journal of Philosophy* 2, 485—493

Sleigh, R. (1964—5), Notes on Chisholm on the Logic of Believing, reprinted in Marras (1972)

Tuomela, R. (1973), Theoretical Concepts in Neobehavioristic Theories, in: Bunge, M., ed., *The Methodological Unity of Science*, Dordrecht, Holland

— (1977), *Human Action and its Explanation*, Dordrecht, Holland